

STIC-Biotech/ChemLib

19/827

From: Whiteman, Brian
Sent: Thursday, November 10, 2003 10:35 AM
To: STIC-Biotech/ChemLib
Subject: seq search

10/696,282 Wilson et al.

SEQ ID NO: 13, 15, and 17 and

nucleotides 335 to 2272 of SEQ ID NO: 1 and nucleotides 1007 to 2272 of SEQ ID NO: 1:

- 1) search against us issued and published us patent database
- 2) commercial databases

thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

RECEIVED
NOV 10 2003
STIC/CHERK Division
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

From: Whiteman, Brian
Sent: Thursday, November 10, 2005 10:41 AM
To: STIC-Biotech/ChemLib
Subject: seq search

10/696,900 Wilson et al.

nucleotides 1 to 143 of SEQ ID NO: 1; nucleotides 4576-4718 of SEQ ID NO: 1; nucleotides 2223-4431 of
SEQ ID NO: 1; *nucleotides* 2829-4432 of SEQ ID NO: 1
1) search against us issued and published us patent database
2) commercial databases

thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

1-4718 NH

11/30/05
JH

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Date completed: _____
Searcher: Beverly e 25 23
Personal time: _____
Elapsed time: _____
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site
____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other CGN

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:24:49 ; Search time 131.527 Seconds
(without alignments)
2458.679 Million cell updates/sec

Title: US-10-696-282-13

Sequence: 1 MAADGYLPDWMEDNLSSEGR.....NNGLYTEPRPIGTRYLTPPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: Genesegp21.*
- 2: genesegp1908.*
- 3: genesegp1908.*
- 4: genesegp2000.*
- 5: genesegp2001.*
- 6: genesegp2002.*
- 7: genesegp2003.*
- 8: genesegp2004.*
- 9: genesegp2005.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3989	100.0	736	3 AAY71167	Aay71167 Adeno-ss
2	3989	100.0	736	6 ABB80232	Abb80232 AAV1 VP1
3	3989	100.0	736	7 ABR62762	AbR62762 Adeno-ss
4	3989	100.0	736	7 ADE76565	AdE76565 Adeno-ss
5	3989	100.0	736	7 ADV70294	AdV70294 Primate a
6	3989	100.0	736	9 ADZ27069	AdZ27069 Adeno-ss
7	3989	100.0	736	9 ADZ27012	AdZ27012 Adeno-ss
8	3985	99.9	735	9 ADV67506	AdV67506 Amino aci
9	3968	99.5	736	9 ADZ27007	AdZ27007 Adeno-ss
10	3963	99.3	736	4 AAB59847	Aab59847 AAV6 caps
11	3963	99.3	736	7 ADE76566	AdE76566 Adeno-ss
12	3963	99.3	736	9 ADV70293	AdV70293 Primate a
13	3963	99.3	736	9 ADZ27070	AdZ27070 Adeno-ss
14	3963	99.3	736	9 ADZ27086	AdZ27086 Adeno-ss
15	3983.5	97.4	737	9 ADZ27010	AdZ27010 Adeno-ss
16	3983	97.3	736	9 ADZ27074	AdZ27074 Adeno-ss
17	3873	97.1	736	9 ADZ27008	AdZ27008 Adeno-ss
18	3853	96.6	736	9 ADZ27009	AdZ27009 Adeno-ss
19	3511	88.0	736	4 AAB59846	Aab59846 AAV3B cap
20	3494	87.6	736	4 AAB59845	Aab59845 AAV3A cap
21	3494	87.6	736	6 ABB80233	Abb80233 AAV3 VP1
22	3494	87.6	736	7 ABR62763	AbR62763 Adeno-ss
23	3494	87.6	736	7 ADE76572	AdE76572 Adeno-ss
24	3494	87.6	736	9 ADV67507	AdV67507 Amino aci

25	3494	87.6	736	9 ADZ27067	AdZ27067 Adeno-ss
26	3483.5	87.3	737	9 ADZ26959	AdZ26959 Adeno-ss
27	3476	87.1	736	9 ADZ27085	AdZ27085 Adeno-ss
28	3473	87.1	736	7 ADE76598	AdE76598 Adeno-ss
29	3469	87.0	738	9 ADZ26960	AdZ26960 Adeno-ss
30	3467	86.9	736	6 ABB80229	Abb80229 AAV9 cap
31	3467	86.9	736	7 ABR62764	AbR62764 Adeno-ss
32	3467	86.9	736	7 ADE76601	AdE76601 Adeno-ss
33	3467	86.9	736	7 ADE76508	AdE76508 Amino aci
34	3464.5	86.9	737	9 ADZ27080	AdZ27080 Adeno-ss
35	3464	86.8	736	7 ADE76597	AdE76597 Adeno-ss
36	3460	86.7	736	7 ADE76599	AdE76599 Adeno-ss
37	3458.5	86.7	737	9 ADZ26963	AdZ26963 Adeno-ss
38	3457	86.7	736	7 ADE76600	AdE76600 Adeno-ss
39	3456.5	86.7	735	9 ADZ27034	AdZ27034 Adeno-ss
40	3456.5	86.7	737	9 ADZ26966	AdZ26966 Adeno-ss
41	3456.5	86.7	737	9 ADZ26962	AdZ26962 Adeno-ss
42	3456	86.6	738	9 ADZ26951	AdZ26951 Adeno-ss
43	3455.5	86.6	737	9 ADZ26965	AdZ26965 Adeno-ss
44	3453.5	86.6	735	9 ADZ26996	AdZ26996 Adeno-ss
45	3453.5	86.6	735	9 ADZ26992	AdZ26992 Adeno-ss

ALIGNMENTS

RESULT 1
AA71167
ID AAY71167 standard; protein; 736 AA.

XX AAY71167;

DT 08-SEP-2000 (first entry)

DE Adeno-associated virus serotype 1 capsid protein VP1.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;

KW cap protein; recombinant viral vector; gene delivery; gene therapy;

KM vaccine; transgene; VP1.

XX Adeno-associated virus 1.

OS Adeno-associated virus 1.

PN WO200028061-A2.

XX 18-MAY-2000.

XX 02-NOV-1999; 99WO-US025694.

XX 05-NOV-1998; 98US-0107114P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Xiao W;

DR WPI, 2000-376571/32.

DR N-PSDB; AAD00772, AAD00777.

PT Novel adeno-associated virus serotype 1 polynucleotide useful for

XX preparation of medicament for delivery of a transgene to a host.

XX Claim 7; Page 87-90; 108pp; English.

PS The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
XX which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40 while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antisense delivery vectors. The AAV-1 does not induce the
CC formation of neutralising antibodies specific to any serotype of AAV
CC hence is useful for transforming host cells, and in the preparation of a

medicament for the delivery of transgene to a host. The present sequence is an AAV-1 cap protein Vp1 which is useful in the production of recombinant viral vector for gene delivery

Sequence 736 AA;

Query Match 100.0%; Score 3989; DB 3; Length 736;
Best Local Similarity 100.0%; Pred. No. 4,9e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MAADGYLPDWLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
Db 1 MAADGYLPDWLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
Oy 61 KGEVNAADAALHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGNTGRAVFO 120
Db 61 KGEVNAADAALHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGNTGRAVFO 120
Oy 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Db 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Oy 181 SVDPDQPLGEBPPATPAVAGPTTMASSGGA PMADNNEGADGVGNASGNHCDSTWLGDRVI 240
Db 181 SVDPDQPLGEBPPATPAVAGPTTMASSGGA PMADNNEGADGVGNASGNHCDSTWLGDRVI 240
Oy 241 TTSPTMALPTYNHNLKYQISSASTGASNDNHRYGYPGFPDNRPHCHFSPPDMQRL 300
Db 241 TTSPTMALPTYNHNLKYQISSASTGASNDNHRYGYPGFPDNRPHCHFSPPDMQRL 300
Oy 301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLSTVQVFSDESYQLPYVLGSAHQ 360
Db 301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLSTVQVFSDESYQLPYVLGSAHQ 360
Oy 361 GCLPPFPADVPMIPOYGYLTLNNGSQAVGRSSFCLEYFPSPQMLRTGNNFTFSYTFEEVP 420
Db 361 GCLPPFPADVPMIPOYGYLTLNNGSQAVGRSSFCLEYFPSPQMLRTGNNFTFSYTFEEVP 420
Oy 421 FHSSYAHQSQSLDRLNPLIDQYLYLNRTQNGSQAOKNDLLFSRGSFAGMSVOPKMWLP 480
Db 421 FHSSYAHQSQSLDRLNPLIDQYLYLNRTQNGSQAOKNDLLFSRGSFAGMSVOPKMWLP 480
Oy 481 GPCYQOQSVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Db 481 GPCYQOQSVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Oy 541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGVHAMG 600
Db 541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGVHAMG 600
Oy 601 ALPGVWQDDRDVYLOGPTIMAKIPHDTGFHPSPLMGFGGLKNPPQIILIKTTPVANPAPA 660
Db 601 ALPGVWQDDRDVYLOGPTIMAKIPHDTGFHPSPLMGFGGLKNPPQIILIKTTPVANPAPA 660
Oy 661 EFSATKFAFSITQYSTGOVSVEIEMELOKENSKRNNPEVQYTSNYSKASANDFTDNNGL 720
Db 661 EFSATKFAFSITQYSTGOVSVEIEMELOKENSKRNNPEVQYTSNYSKASANDFTDNNGL 720
Oy 721 YTEPRPIGRTYLRPL 736
Db 721 YTEPRPIGRTYLRPL 736

```

RESULT 2
ABB80232
ID ABB80232 standard; protein; 736 AA.

XX ABB80232;
XX AC
XX 20-NOV-2003 (first entry)
XX AAV1 vp1 protein.
XX

KW Adeno-associated virus; AAV; serotype 9; rep; cap; vp1; vp2; vp3;
KW splice variant; transgene.

OS Adeno associated virus serotype 1.

XX WO2003052052-A2.

XX 26-JUN-2003.

XX 12-NOV-2002; 2002WO-US033631.

XX 17-DEC-2001; 2001US-0341150P.

XX 05-JUN-2002; 2002US-0386132P.

XX (TYPE-) UNIV PENNSYLVANIA.

XX Gao G, Wilson JM, Alvira M;

XX WPI; 2003-523523/49.

XX New isolated adeno-associated virus (AAV) comprising an AAV capsid.

XX Useful for preparing a medicament for delivering a transgene to a cell.

XX Disclosure; Fig 2; 76pp; English.

The sequences given in ABB80231-34 represent vp1 proteins derived from various adeno-associated virus (AAV) serotypes. These sequences were used in the scope of the invention for comparison with the cap protein derived from AAV serotype 9. The AAV capsid comprises three protein vp1, vp2 and vp3, which are alternative splice variants. The AAV or the nucleic acid molecule is useful for preparing a medicament for delivering a transgene to a cell

XX Sequence 736 AA;

Query Match 100.0%; Score 3989; DB 6; Length 736;
Best Local Similarity 100.0%; Pred. No. 4,9e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MAADGYLPDWLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
Db 1 MAADGYLPDWLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
Oy 61 KGEVNAADAALHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGNTGRAVFO 120
Db 61 KGEVNAADAALHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGNTGRAVFO 120
Oy 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Db 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Oy 181 SVDPDQPLGEBPPATPAVAGPTTMASSGGA PMADNNEGADGVGNASGNHCDSTWLGDRVI 240
Db 181 SVDPDQPLGEBPPATPAVAGPTTMASSGGA PMADNNEGADGVGNASGNHCDSTWLGDRVI 240
Oy 241 TTSPTMALPTYNHNLKYQISSASTGASNDNHRYGYPGFPDNRPHCHFSPPDMQRL 300
Db 241 TTSPTMALPTYNHNLKYQISSASTGASNDNHRYGYPGFPDNRPHCHFSPPDMQRL 300
Oy 301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLSTVQVFSDESYQLPYVLGSAHQ 360
Db 301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLSTVQVFSDESYQLPYVLGSAHQ 360
Oy 361 GCLPPFPADVPMIPOYGYLTLNNGSQAVGRSSFCLEYFPSPQMLRTGNNFTFSYTFEEVP 420
Db 361 GCLPPFPADVPMIPOYGYLTLNNGSQAVGRSSFCLEYFPSPQMLRTGNNFTFSYTFEEVP 420
Oy 421 FHSSYAHQSQSLDRLNPLIDQYLYLNRTQNGSQAOKNDLLFSRGSFAGMSVOPKMWLP 480
Db 421 FHSSYAHQSQSLDRLNPLIDQYLYLNRTQNGSQAOKNDLLFSRGSFAGMSVOPKMWLP 480
Oy 481 GPCYQOQSVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540

```


OS Unidentified.
 XX EPI310571-A2.
 XX
 PD 14-MAY-2003.
 XX
 PF 12-NOV-2002; 2002EP-00257826.
 XX
 PR 13-NOV-2001; 2001US-0350607P.
 PR 17-DEC-2001; 2001US-0341117P.
 PR 01-MAY-2002; 2002US-0377066P.
 PR 05-JUN-2002; 2002US-0386675P.
 XX
 PA (UNPE-) UNIV PENNSYLVANIA.
 XX
 PI Gao G, Wilson JM, Alvira M;
 DR WPI; 2003-450984/43.
 XX
 PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
 PT preventing or treating hyperproliferative or autoimmune diseases,
 PT comprising subjecting a sample having a DNA to amplification via
 PT polymerase chain reaction.
 XX
 PS Disclosure; SEQ ID NO 64; 419pp; English.
 XX
 CC The invention relates to a novel method for detecting adeno-associated
 CC virus (AAV) sequences in a sample, which comprises subjecting a sample
 CC containing a DNA to amplification via a polymerase chain reaction (PCR).
 CC The AAV sequence have the following activities: cytostatic,
 CC antiproliferative, antineoplastic, antitumor, neuroprotective,
 CC antidiabetic, antihypertensive, dermatological, and antiinflammatory. The AAV
 CC sequence can be used in gene therapy or as part of a vaccine to treat
 CC disorders. The method is useful in detecting and/or identifying AAV
 CC sequences and isolating novel sequences that are identified. The
 CC sequences may be used e.g. for preventing or treating hyperproliferative
 CC conditions such as cancers and psoriasis, and other autoimmune diseases
 CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
 CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
 CC AAV related protein sequence of the invention.
 XX
 XX Sequence 736 AA;
 SQ
 Query Match 100.0%; Score 3989; DB 7; Length 736;
 Best Local Similarity 100.0%; Pred. No. 4.9e-311;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAADGYLPDWEEDNLSEGIKREWDLKPAPKPKANQKQDDGRGLVLPQYKYLGFPGSLD 60
 DB 1 MAADGYLPDWEEDNLSEGIKREWDLKPAPKPKANQKQDDGRGLVLPQYKYLGFPGSLD 60
 QY 61 KCEPNAADAALHDKAYDQDLKAGDNPYLRYNADAFQORLOEDPSPFGNLCRAAFQ 120
 DB 61 KCEPNAADAALHDKAYDQDLKAGDNPYLRYNADAFQORLOEDPSPFGNLCRAAFQ 120
 QY 121 AKKRVLEPLGLVEBGAKTPAGKRPVEQSPQEPDSSSGIGKYGQOPAKKRLNFGQTGSE 180
 DB 121 AKKRVLEPLGLVEBGAKTPAGKRPVEQSPQEPDSSSGIGKYGQOPAKKRLNFGQTGSE 180
 QY 181 SVDPDQPLGEBPPATPAVGPPTTMASSGGA PMADNNEGADGVNAGNMHCSTWLGDRVI 240
 DB 181 SVDPDQPLGEBPPATPAVGPPTTMASSGGA PMADNNEGADGVNAGNMHCSTWLGDRVI 240
 QY 241 TTSTRTMALPTYNHLYKQISASTGASNDNHYFGYSTPWGFDNRRPHCHSPDQRL 300
 DB 241 TTSTRTMALPTYNHLYKQISASTGASNDNHYFGYSTPWGFDNRRPHCHSPDQRL 300
 QY 301 INNNGFPRKRLNFKLNFQVKEVYTTNDGVTTIANNLSTVQVPSDSEYQLPYVLSAHO 360
 DB 301 INNNGFPRKRLNFKLNFQVKEVYTTNDGVTTIANNLSTVQVPSDSEYQLPYVLSAHO 360
 QY 361 GCLPPFPADVFMIPQYGYLTLLNGSOAVGRSSFYCLEYPPSQMLRTGNNTFSYTEBEVP 420

DB 361 GCLPPFPADVFMIPQYGYLTLLNGSOAVGRSSFYCLEYPPSQMLRTGNNTFSYTEBEVP 420
 QY 421 FHSSYASQSLDRMLNPLIDQVLYLNRTONOGSAGQNKDLFSRGS PAGMSYQPKNMLP 480
 DB 421 FHSSYASQSLDRMLNPLIDQVLYLNRTONOGSAGQNKDLFSRGS PAGMSYQPKNMLP 480
 QY 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
 DB 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
 QY 541 MIRCKESAGASNTALDVMITDDEEIKATNPVATERGTAVNVPQSSSTDPATGDVAMG 600
 DB 541 MIRCKESAGASNTALDVMITDDEEIKATNPVATERGTAVNVPQSSSTDPATGDVAMG 600
 QY 601 ALPGMWQDQDVYLQGPIMAKIPHTDGHFSPPLMGFGKKNPPQILIKNTVPANPPA 660
 DB 601 ALPGMWQDQDVYLQGPIMAKIPHTDGHFSPPLMGFGKKNPPQILIKNTVPANPPA 660
 QY 661 EFSATKFASTITQYSTQVSEIEMELQKNSKRMNDEVOYTSNYSKAVDFTVNNGL 720
 DB 661 EFSATKFASTITQYSTQVSEIEMELQKNSKRMNDEVOYTSNYSKAVDFTVNNGL 720
 QY 721 YTEPRPIGTRYLRPL 736
 DB 721 YTEPRPIGTRYLRPL 736
 RESULT 5
 ID ADV70294
 ADV70294 standard; protein; 736 AA.
 XX
 XX ADV70294;
 DT 10-MAR-2005 (first entry)
 XX
 DE Primate adeno-associated virus 1 capsid protein VP1.
 XX
 XX immunosuppressive; gene therapy; immunogenicity; virus inactivation;
 XX hemophilia; Pepck deficiency; galactosemia; phenylketonuria;
 KW Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
 KW atherosclerosis; thrombosis; embolism; Parkinson's disease;
 KW congestive heart failure; cancer; inflammation; immune disorder;
 KW muscular dystrophy; diabetes; VPI.
 XX
 OS Adeno-associated virus 1.
 XX
 XX WO2004112727-A2.
 XX
 PD 29-DEC-2004.
 XX
 PF 21-JUN-2004; 2004WO-US019884.
 XX
 PR 19-JUN-2003; 2003US-0480395P.
 PR 30-APR-2004; 2004US-0567310P.
 PR 03-JUN-2004; 2004US-0576501P.
 XX
 PA (AVIG-) AVIGEN INC.
 XX
 PI Arbetman AE, Colosi P, Lochrie MA, Surosky RT;
 DR WPI; 2005-048755/05.
 XX
 PT New mutated adeno-associated virus (AAV) capsid protein that when present
 PT in an AAV virion imparts decreased immunoreactivity to the virion as
 PT compared to the corresponding wild-type virion, useful for treating e.g.
 PT hemophilia.
 XX
 PS Example 5; SEQ ID NO 20; 136pp; English.
 XX
 CC The invention describes a mutated adeno-associated virus (AAV) capsid
 CC protein that when present in an AAV virion imparts decreased
 CC immunoreactivity to the virion as compared to the corresponding wild-type
 CC virion. Also described are: a polynucleotide encoding the mutated protein

CC above; a recombinant AAV virion comprising the mutated protein above; and
CC delivering a recombinant AAV virion to a cell or tissue of a vertebrate
CC subject. The recombinant AAV virion is useful for delivering a
CC heterologous nucleic acid molecule to a cell or tissue of a vertebrate
CC subject, where the protein encoded by the heterologous nucleic acid
CC molecule is expressed at a level that provides a therapeutic effect.
CC where the recombinant AAV virion may comprise a non-primate, mammalian
CC AAV capsid protein that when present in an AAV virion imparts decreased
CC immunoreactivity to the virion as compared to immunoreactivity of primate
CC AAV-2, and the heterologous nucleic acid molecule, where the heterologous
CC nucleic acid molecule encodes a therapeutic protein and is operably
CC linked to control elements capable of directing the in vivo transcription
CC and translation of the protein. The protein or the recombinant AAV virion
CC is useful for treating or preventing a wide variety of disorders such as
CC hemophilia, glycogen storage deficiency type 1a, Pepck deficiency,
CC galactosemia, phenylketonuria, Crigler-Najjar disease, Gout and Lesch-
CC Nyan syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,
CC thrombosis, embolism, Parkinson's disease, congestive heart failure,
CC cancer, inflammatory and immune disorders, muscular dystrophies, and
CC diabetes. This is the amino acid sequence of adeno-associated virus 1
CC (AAV1) capsid protein Vp1.
CC
CC
XX

Sequence 736 AA:

Query Match 100.0%; Score 3989; DB 9; Length 736;
Best Local Similarity 100.0%; Pred. No. 4,9e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGTLPLWLENDNLSEGIREWMDLKPQAPKPKANQKODDGGGLVLPQYKYLGPENGID 60
DB 1 MAADGTLPLWLENDNLSEGIREWMDLKPQAPKPKANQKODDGGGLVLPQYKYLGPENGID 60
QY 61 KGEPVAAADAAALEHKAAYDQOLKAGDNPLYRNHADAFOERLQEDTSFGNLAGRAVQ 120
DB 61 KGEPVAAADAAALEHKAAYDQOLKAGDNPLYRNHADAFOERLQEDTSFGNLAGRAVQ 120
QY 121 AKKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKTQOPAKKRLNFGQTGDS 180
DB 121 AKKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKTQOPAKKRLNFGQTGDS 180
QY 121 AKKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKTQOPAKKRLNFGQTGDS 180
DB 121 AKKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKTQOPAKKRLNFGQTGDS 180
QY 181 SVDPDPLGEPPTTPAAGPTTMASSGAGPMAADNNGAGCGVAGSAGNMHCDSTWLDRLY 240
DB 181 SVDPDPLGEPPTTPAAGPTTMASSGAGPMAADNNGAGCGVAGSAGNMHCDSTWLDRLY 240
QY 181 SVDPDPLGEPPTTPAAGPTTMASSGAGPMAADNNGAGCGVAGSAGNMHCDSTWLDRLY 240
DB 181 SVDPDPLGEPPTTPAAGPTTMASSGAGPMAADNNGAGCGVAGSAGNMHCDSTWLDRLY 240
QY 241 TTSRTMALPTYNNHLYKOISASTGASNDNHFGYSTPMGYDFFRHHCFSPRDMQRL 300
DB 241 TTSRTMALPTYNNHLYKOISASTGASNDNHFGYSTPMGYDFFRHHCFSPRDMQRL 300
QY 241 TTSRTMALPTYNNHLYKOISASTGASNDNHFGYSTPMGYDFFRHHCFSPRDMQRL 300
DB 241 TTSRTMALPTYNNHLYKOISASTGASNDNHFGYSTPMGYDFFRHHCFSPRDMQRL 300
QY 301 INNMGFRPKRLNFKLNIQKEVTNDGVTTIANLSTVQVFSDEYQLPYVLGSAHQ 360
DB 301 INNMGFRPKRLNFKLNIQKEVTNDGVTTIANLSTVQVFSDEYQLPYVLGSAHQ 360
QY 301 INNMGFRPKRLNFKLNIQKEVTNDGVTTIANLSTVQVFSDEYQLPYVLGSAHQ 360
DB 301 INNMGFRPKRLNFKLNIQKEVTNDGVTTIANLSTVQVFSDEYQLPYVLGSAHQ 360
QY 361 GCLPPPADVFMIPOYGYLTLLNGSQAVGSSFFCYCLEFPSSQMLRTGNNFTFYTEEVP 420
DB 361 GCLPPPADVFMIPOYGYLTLLNGSQAVGSSFFCYCLEFPSSQMLRTGNNFTFYTEEVP 420
QY 361 GCLPPPADVFMIPOYGYLTLLNGSQAVGSSFFCYCLEFPSSQMLRTGNNFTFYTEEVP 420
DB 361 GCLPPPADVFMIPOYGYLTLLNGSQAVGSSFFCYCLEFPSSQMLRTGNNFTFYTEEVP 420
QY 421 FHSSYASHQSGLDLRMLPPLIDQVLYYLRNTQNGSAGNKKDLRSRGSPPAGMSYQPKMWLP 480
DB 421 FHSSYASHQSGLDLRMLPPLIDQVLYYLRNTQNGSAGNKKDLRSRGSPPAGMSYQPKMWLP 480
QY 421 FHSSYASHQSGLDLRMLPPLIDQVLYYLRNTQNGSAGNKKDLRSRGSPPAGMSYQPKMWLP 480
DB 421 FHSSYASHQSGLDLRMLPPLIDQVLYYLRNTQNGSAGNKKDLRSRGSPPAGMSYQPKMWLP 480
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLANGESIIINFGTAMASHKDEDEKFFPMGCV 540
DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLANGESIIINFGTAMASHKDEDEKFFPMGCV 540
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLANGESIIINFGTAMASHKDEDEKFFPMGCV 540
DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLANGESIIINFGTAMASHKDEDEKFFPMGCV 540
QY 541 MIRGKESAGASNTALDNVMTTDEBEIKATNPVATERGTVAVVPOSSSTPAGCDVAMG 600
DB 541 MIRGKESAGASNTALDNVMTTDEBEIKATNPVATERGTVAVVPOSSSTPAGCDVAMG 600
QY 541 MIRGKESAGASNTALDNVMTTDEBEIKATNPVATERGTVAVVPOSSSTPAGCDVAMG 600
DB 541 MIRGKESAGASNTALDNVMTTDEBEIKATNPVATERGTVAVVPOSSSTPAGCDVAMG 600
QY 601 ALPQWVQODRDVYLQGPIMAKIPHTDGHFSPPLMGFGFKNPPQLIKNTTVPANPPA 660
DB 601 ALPQWVQODRDVYLQGPIMAKIPHTDGHFSPPLMGFGFKNPPQLIKNTTVPANPPA 660
QY 601 ALPQWVQODRDVYLQGPIMAKIPHTDGHFSPPLMGFGFKNPPQLIKNTTVPANPPA 660
DB 601 ALPQWVQODRDVYLQGPIMAKIPHTDGHFSPPLMGFGFKNPPQLIKNTTVPANPPA 660
QY 661 EFSATKFASTITQYSTQGVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNGL 720

DB 661 EFSATKFASTITQYSTQGVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNGL 720
QY 721 YTEPRPICTRYLTRPL 736
DB 721 YTEPRPICTRYLTRPL 736

RESULT 6

AD227069 ID AD227069 standard; protein; 736 AA.

AD227069; 30-JUN-2005 (first entry)

Adeno-associated virus protein SEQ ID NO 219.

KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM bacterial infection; cancer; ulcerative colitis; antineumatic;
KM antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KM antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM viricide; antibacterial; cyostatic; antitumor; dermatological.

Adeno-associated virus.

W020503321-A2.

14-APR-2005.

30-SEP-2004; 2004MO-US028817.

30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.

(TYPE-) UNIV PENNSYLVANIA.

Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

WPI; 2005-285437/29.

PT New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritides, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.

Disclosure; SEQ ID NO 219; 569pp; English.

CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.

Sequence 736 AA:

Query Match 100.0%; Score 3989; DB 9; Length 736;
Best Local Similarity 100.0%; Pred. No. 4,9e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGTLPLWLENDNLSEGIREWMDLKPQAPKPKANQKODDGGGLVLPQYKYLGPENGID 60
DB 1 MAADGTLPLWLENDNLSEGIREWMDLKPQAPKPKANQKODDGGGLVLPQYKYLGPENGID 60
QY 61 KGEPVAAADAAALEHKAAYDQOLKAGDNPLYRNHADAFOERLQEDTSFGNLAGRAVQ 120

Db 61 KGEVNAADAAALEHDKAYDQOLKAGDNPLYLRYNHADAFOERLQEDTSPFCGNLGRAVFO 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKKRVEQSPQEPSSSGIGTGOQPAKKRLNFCQTDGSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKKRVEQSPQEPSSSGIGTGOQPAKKRLNFCQTDGSE 180
QY 181 SVDPQPLGEPBPATPAVAGPTTMAAGGAPMADNNEGADVGNAAGNMHCSTWLGDRVI 240
Db 181 SVDPQPLGEPBPATPAVAGPTTMAAGGAPMADNNEGADVGNAAGNMHCSTWLGDRVI 240
QY 241 TTSRTMALPTYYNNHLYKQISSASTGASNDNHYFGYSTPMGWFDNRFHCHFSRDMQRL 300
Db 241 TTSRTMALPTYYNNHLYKQISSASTGASNDNHYFGYSTPMGWFDNRFHCHFSRDMQRL 300
QY 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTITANNLTSTVQVPSDSBYQLPYVLGSAHQ 360
Db 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTITANNLTSTVQVPSDSBYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVPMIPOYGYLTLLNGSQAAGRSFPCLEYFPSPQMLRTGNFTFSTFEFVP 420
Db 361 GCLPPFPADVPMIPOYGYLTLLNGSQAAGRSFPCLEYFPSPQMLRTGNFTFSTFEFVP 420
QY 421 FHSSTYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDLIFSRGSPAGMSVOPKMWLP 480
Db 421 FHSSTYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDLIFSRGSPAGMSVOPKMWLP 480
QY 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYKLNGBRESIINPGTAMASHKDEDKFFPMGCV 540
Db 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYKLNGBRESIINPGTAMASHKDEDKFFPMGCV 540
QY 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPAVDVHAMG 600
Db 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPAVDVHAMG 600
QY 601 ALPGVWQDRDYYLQGPYIAKIPHTDGHFHPSPMLGSGGLKNPPQIILIKTVPVAPNPA 660
Db 601 ALPGVWQDRDYYLQGPYIAKIPHTDGHFHPSPMLGSGGLKNPPQIILIKTVPVAPNPA 660
QY 661 FESATKFASTFQYSTGOVSVEIEMELQENSKRNNPEVQYTSNYSKASANDFTVDNGL 720
Db 661 FESATKFASTFQYSTGOVSVEIEMELQENSKRNNPEVQYTSNYSKASANDFTVDNGL 720
QY 721 YTEPRPIGTRVLTREPL 736
Db 721 YTEPRPIGTRVLTREPL 736
RESULT 7
AD227012 standard; protein; 736 AA.
ID AD227012
XX
AC AD227012;
XX
XX 30-JUN-2005 (first entry)
DE Adeno-associated virus protein SEQ ID NO 162.
XX
XX
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antithrombotic;
KW antitubercular; neuroprotective; antiinflammatory; antidiabetic;
KW antiparasitic; vasotropic; gastrointestinal; gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS Adeno-associated virus.
XX
XX WO200503321-A2.
XX
XX 14-APR-2005.
PD
XX
XX 30-SEP-2004; 2004WO-US028817.
PF
XX
XX 30-SEP-2003; 2003US-0508226P.
PR

PR 29-APR-2004; 2004US-0566546P.
XX
XX (UPE-) UNIV PENNSYLVANIA.
PA
XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
PI
XX WPI; 2005-285437/29.
DR
XX
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX
PS Disclosure; SEQ ID NO 162; 569bp; English.
XX
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
SQ Sequence 736 AA;

Query Match 100.0%; Score 3989; DB 9; Length 736;
Best local similarity 100.0%; Pred. No. 4, 9e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGCYLPMLENDLSEGIREWMDLKGAPKPKANQKODRGGLVPGYKYLGPFGGLD 60
Db 1 MAAGCYLPMLENDLSEGIREWMDLKGAPKPKANQKODRGGLVPGYKYLGPFGGLD 60
QY 61 KGEVNAADAAALEHDKAYDQOLKAGDNPLYLRYNHADAFOERLQEDTSPFCGNLGRAVFO 120
Db 61 KGEVNAADAAALEHDKAYDQOLKAGDNPLYLRYNHADAFOERLQEDTSPFCGNLGRAVFO 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKKRVEQSPQEPSSSGIGTGOQPAKKRLNFCQTDGSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKKRVEQSPQEPSSSGIGTGOQPAKKRLNFCQTDGSE 180
QY 181 SVDPQPLGEPBPATPAVAGPTTMAAGGAPMADNNEGADVGNAAGNMHCSTWLGDRVI 240
Db 181 SVDPQPLGEPBPATPAVAGPTTMAAGGAPMADNNEGADVGNAAGNMHCSTWLGDRVI 240
QY 241 TTSRTMALPTYYNNHLYKQISSASTGASNDNHYFGYSTPMGWFDNRFHCHFSRDMQRL 300
Db 241 TTSRTMALPTYYNNHLYKQISSASTGASNDNHYFGYSTPMGWFDNRFHCHFSRDMQRL 300
QY 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTITANNLTSTVQVPSDSBYQLPYVLGSAHQ 360
Db 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTITANNLTSTVQVPSDSBYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVPMIPOYGYLTLLNGSQAAGRSFPCLEYFPSPQMLRTGNFTFSTFEFVP 420
Db 361 GCLPPFPADVPMIPOYGYLTLLNGSQAAGRSFPCLEYFPSPQMLRTGNFTFSTFEFVP 420
QY 421 FHSSTYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDLIFSRGSPAGMSVOPKMWLP 480
Db 421 FHSSTYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDLIFSRGSPAGMSVOPKMWLP 480
QY 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYKLNGBRESIINPGTAMASHKDEDKFFPMGCV 540
Db 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYKLNGBRESIINPGTAMASHKDEDKFFPMGCV 540
QY 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPAVDVHAMG 600
Db 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPAVDVHAMG 600


```

QY      601 ALPGWVMDRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGELKNPPOILLKNTPVANPPA 660
      |||
DB      601 ALPGWVMDRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGELKNPPOILLKNTPVANPPA 660
QY      661 EFSATKFAFIQYSTGVSVSEIEMELQENSKRNKPNPEVQYTSNYSKASANVDFVDNNGI 720
      |||
DB      661 EFSATKFAFIQYSTGVSVSEIEMELQENSKRNKPNPEVQYTSNYSKASANVDFVDNNGI 720
QY      721 YTEPRPIGTRVYLTRPL 736
      |||
DB      721 YTEPRPIGTRVYLTRPL 736

RESULT 8
ADV67506
ID      ADV67506 standard; protein; 735 AA.
XX
AC      ADV67506;
XX
DT      10-MAR-2005 (first entry)
XX
DE      Amino acid sequence of the capsid protein of AAV serotype 1.
XX
KM      antiherpesoviral; antilipemic; gene therapy; cholesterol;
XX      apolipoprotein E; apoE; apolipoprotein A; apoA; atherosclerosis;
XX      lipoprotein defect; capsid protein.
OS      Adeno-associated virus.
XX
PN      MO2004108922-A2.
XX
PD      16-DEC-2004.
XX
PF      23-APR-2004; 2004MO-US010965.
XX
PR      25-APR-2003; 2003US-0465293P.
XX
PA      (UYPE-) UNIV PENNSYLVANIA.
XX
PI      Rader DJ, Wilson JM;
XX
DR      WPI; 2005-031700/03.
XX
PT      Lowering total cholesterol levels and treating atherosclerosis in a
XX      subject comprises delivering a recombinant adeno-associated virus (AAV)
XX      comprising an AAV serotype capsid protein or a gene encoding human
XX      apolipoprotein E (apoE) or apoA.
XX
PS      Disclosure; SEQ ID NO 4; 69pp; English.
XX
XX
CC      The specification describes a method for lowering total cholesterol
CC      levels in a subject. The method comprises delivering to the subject a
CC      recombinant adeno-associated virus (AAV) comprising a gene encoding a
CC      human apolipoprotein E (apoE) or apoA under the control of a regulatory
CC      control sequence which directs expression of the gene. The recombinant
CC      AAV also comprises a capsid protein selected from an AAV serotype (e.g.
CC      AAV7 or AAV8) which preferentially expresses high levels of transgene in
CC      liver. A therapeutically effective amount of apoE or apoA expression is
CC      obtained upon delivery of low dose of AAV. The method of the invention is
CC      useful for lowering total cholesterol levels in a subject, e.g. for
CC      treating atherosclerosis, and for correcting defects in lipoprotein. The
CC      present sequence represents a capsid protein of AAV serotype 1 (AAV1),
CC      which may be used in recombinant AAV vectors of the invention.
XX
SQ      Sequence 735 AA;

Query Match      99.9%; Score 3985; DB 9; Length 735;
Best Local Similarity 100.0%; Pred. No. 1e-310;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MAADGYLPDWLENDLSGIRFEMWDLKGAAPKPKANOOKODDGRGLVPGKYLGPFNGLD 60
      |||
DB      1 MAADGYLPDWLENDLSGIRFEMWDLKGAAPKPKANOOKODDGRGLVPGKYLGPFNGLD 60

```

```

QY      61 KGEPVNAADAAALBHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSFGNLGRAVFO 120
      |||
DB      61 KGEPVNAADAAALBHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSFGNLGRAVFO 120
QY      121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSGIGKTGOQPAKKRLNFGQTGSE 180
      |||
DB      121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSGIGKTGOQPAKKRLNFGQTGSE 180
QY      181 SVDPDQPLGEPPTATPAVAGPTTMAAGGAPADNNEGADGVNAGNHCSTWLGDRYI 240
      |||
DB      181 SVDPDQPLGEPPTATPAVAGPTTMAAGGAPADNNEGADGVNAGNHCSTWLGDRYI 240
QY      241 TTSTRTWALPTYNHNLVYQOISSASTGASNDNHYFGYSTPMGVFEDNRFHCHSPDMOQL 300
      |||
DB      241 TTSTRTWALPTYNHNLVYQOISSASTGASNDNHYFGYSTPMGVFEDNRFHCHSPDMOQL 300
QY      301 INNNMGFRPKRLNPLFNIQVKEVTTNDGVTTIANNLSTVQVFPDSEYQLPYVLSAHQ 360
      |||
DB      301 INNNMGFRPKRLNPLFNIQVKEVTTNDGVTTIANNLSTVQVFPDSEYQLPYVLSAHQ 360
QY      361 GCLPFPADVFEMIPOGYVLTLLNGSQAVGRSSFYCLEYFPSCMLRTGNNFTSYTPEEVP 420
      |||
DB      361 GCLPFPADVFEMIPOGYVLTLLNGSQAVGRSSFYCLEYFPSCMLRTGNNFTSYTPEEVP 420
QY      421 FHSSTAHQSGLDRLMNPILIDYLYLNTTQNGSGANQNDLLFSRGSPPAGMSVOPKNWLP 480
      |||
DB      421 FHSSTAHQSGLDRLMNPILIDYLYLNTTQNGSGANQNDLLFSRGSPPAGMSVOPKNWLP 480
QY      481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLGRSIIINPGTAMASHKDEDEKFFPMSCV 540
      |||
DB      481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLGRSIIINPGTAMASHKDEDEKFFPMSCV 540
QY      541 MIFGKESAGASNTALDNWMTDEBEIKATNPATERFGTAVNFOSSSTDPAIGDVHANG 600
      |||
DB      541 MIFGKESAGASNTALDNWMTDEBEIKATNPATERFGTAVNFOSSSTDPAIGDVHANG 600
QY      601 ALPGWVMDRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGELKNPPOILLKNTPVANPPA 660
      |||
DB      601 ALPGWVMDRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGELKNPPOILLKNTPVANPPA 660
QY      661 EFSATKFAFIQYSTGVSVSEIEMELQENSKRNKPNPEVQYTSNYSKASANVDFVDNNGI 720
      |||
DB      661 EFSATKFAFIQYSTGVSVSEIEMELQENSKRNKPNPEVQYTSNYSKASANVDFVDNNGI 720
QY      721 YTEPRPIGTRVYLTRP 735
      |||
DB      721 YTEPRPIGTRVYLTRP 735

RESULT 9
AD227007
ID      AD227007 standard; protein; 736 AA.
XX
AC      AD227007;
XX
DT      30-JUN-2005 (first entry)
XX
DE      Adeno-associated virus protein SEQ ID NO 157.
XX
XX
KM      rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM      scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM      bacterial infection; cancer; ulcerative colitis; antineoplastic;
KM      antidiabetic; neuroprotective; antiinflammatory; antidiabetic;
KM      antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM      virucide; antibacterial; cytostatic; antulcer; dermatological.
XX
XX
OS      Adeno-associated virus.
XX
PN      MO2005033321-A2.
XX
PD      14-APR-2005.
XX

```

```

PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-056546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI, 2005-285437/29.
DR
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
XX Claim 19, SEQ ID NO 157; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
SQ Sequence 736 AA;
Query Match 99.5%; Score 3968; DB 9; Length 736;
Beet Local Similarity 99.3%; Pred. No. 2.4e-309;
Matches 731; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAADGYLPDWMLENDLSEGIKREWMDLKGAPKPKANQOKDDGRGLVPGYKYLGPFGNGLD 60
DB 1 MAADGYLPDWMLENDLSEGIKREWMDLKGAPKPKANQOKDDGRGLVPGYKYLGPFGNGLD 60
QY 61 KGEPVNADAAALBHDKAYDQOLKAGDNPYLRYNHADAFOERLOEDTSPGNGLGRAVFO 120
DB 61 KGEPVNADAAALBHDKAYDQOLKAGDNPYLRYNHADAFOERLOEDTSPGNGLGRAVFO 120
QY 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEOSPOEDSSSGIGKGOQPAKKRLNFGQTGSE 180
DB 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEOSPOEDSSSGIGKGOQPAKKRLNFGQTGSE 180
QY 181 SVDPDQPLGEPPTAAVPTTMAAGGAPMADNNEGADGVNAGSNHCHSTWLGDRVI 240
DB 181 SVDPDQPLGEPPTAAVPTTMAAGGAPMADNNEGADGVNAGSNHCHSTWLGDRVI 240
QY 241 TTSTRTMALPTVNNHLYKOISSASTGASNDNHYPGYSTPMWGYDFNRRPHCHSPDWQRL 300
DB 241 TTSTRTMALPTVNNHLYKOISSASTGASNDNHYPGYSTPMWGYDFNRRPHCHSPDWQRL 300
QY 301 INNMGFRPKRLNFKLNFNIOVEKVTNNGVTTIANNLSTVOVPSDSEYQLPYVIGSAHQ 360
DB 301 INNMGFRPKRLNFKLNFNIOVEKVTNNGVTTIANNLSTVOVPSDSEYQLPYVIGSAHQ 360
QY 361 GCLPFPADVPMIPOYGYLTLLNGSOAVGRSSFCLEYPSOMLRTGNFTSYFEEVP 420
DB 361 GCLPFPADVPMIPOYGYLTLLNGSOAVGRSSFCLEYPSOMLRTGNFTSYFEEVP 420
QY 421 FHSSVAHSGSLDRLNPLIDQYLYLNRTQNGSGAQNKKDLLPSRGSPPAGMSVQPKANLP 480
DB 421 FHSSVAHSGSLDRLNPLIDQYLYLNRTQNGSGAQNKKDLLPSRGSPPAGMSVQPKANLP 480
QY 481 GGCYQOQRSKTKTDNNNSNFTWTGASKYNLNGRBSITNPGTAMASHDDEKFFPMGCV 540
DB 481 GGCYQOQRSKTKTDNNNSNFTWTGASKYNLNGRBSITNPGTAMASHDDEKFFPMGCV 540
QY 541 MIFGKESAGASNTALDNNMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600

```

```

DB 541 MIFGKESAGASNTALDNNMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
QY 601 ALPGMWODDDVYLYQGIWAKIPHTDGHFHPSPLMGFGKKNPPOLIKNTPPANPPA 660
DB 601 ALPGMWODDDVYLYQGIWAKIPHTDGHFHPSPLMGFGKKNPPOLIKNTPPANPPA 660
QY 661 EFSATKRFASFTQYSTQGVSEIEMELQKNSKRMNDEVQYTSNYSKASAVDFTVDNNGL 720
DB 661 EFSATKRFASFTQYSTQGVSEIEMELQKNSKRMNDEVQYTSNYSKASAVDFTVDNNGL 720
QY 721 YTEBRPIGTRYLTRPL 736
DB 721 YTEBRPIGTRYLTRPL 736
RESULT 10
AAB59847
ID AAB59847 standard; protein; 736 AA.
XX
AC AAB59847;
XX
DT 28-MAR-2001 (first entry)
XX
DE AAV6 capsid protein VP1.
XX
KW AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; cancer; acquired immunodeficiency
KW blood clotting disorder; diabetes; capsid protein VP1.
XX
OS Adeno associated virus.
XX
PN US6156303-A.
XX
PD 05-DEC-2000.
XX
PF 11-JUN-1997; 97US-00873168.
XX
PR 11-JUN-1997; 97US-00873168.
XX
PA (UNIW ) UNIV WASHINGTON.
XX
PI Russell DW, Rutledge EA;
XX
DR WPI, 2001-060164/07.
XX
PT Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes.
XX
PS Disclosure; Fig 2; 50pp; English.
XX
CC The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein VP1 of one such serotype (AAV6). AAV6
CC can be used to construct AAV viral vectors for use in gene therapy for a
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell
XX
SQ Sequence 736 AA;
Query Match 99.3%; Score 3963; DB 4; Length 736;
Beet Local Similarity 99.2%; Pred. No. 6.1e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAADGYLPDWMLENDLSEGIKREWMDLKGAPKPKANQOKDDGRGLVPGYKYLGPFGNGLD 60
DB 1 MAADGYLPDWMLENDLSEGIKREWMDLKGAPKPKANQOKDDGRGLVPGYKYLGPFGNGLD 60
QY 61 KGEPVNADAAALBHDKAYDQOLKAGDNPYLRYNHADAFOERLOEDTSPGNGLGRAVFO 120

```



```

Db      |||
61      KGEVNAADAALEHDKAYDQOLKAGDNPYLRYNADAEFORLEEDTSFGGNLGRAVFO 120
Qy      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEOSPOBPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Db      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEOSPOBPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Qy      181 SVPDPQPLGEPPTATPAVGPPTMASGGGAPMADNNEGADGVGNAGNHCDSITWLGDRVI 240
Db      181 SVPDPQPLGEPPTATPAVGPPTMASGGGAPMADNNEGADGVGNAGNHCDSITWLGDRVI 240
Qy      241 TTSTRTMALPTYNNHLYKQISSASTGASNDNHFGYSTWGWGFDNRFCHSPRMDORL 300
Db      241 TTSTRTMALPTYNNHLYKQISSASTGASNDNHFGYSTWGWGFDNRFCHSPRMDORL 300
Qy      301 INNMGFPKRLNFKLFNIOVKEVTTNDGVTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
Db      301 INNMGFPKRLNFKLFNIOVKEVTTNDGVTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
Qy      361 GCLPPFPADVPMIPOYGYLTLLNGSQAVERSSFYCLEYPPSQMLRTGNFTFSYTFEYV 420
Db      361 GCLPPFPADVPMIPOYGYLTLLNGSQAVERSSFYCLEYPPSQMLRTGNFTFSYTFEYV 420
Qy      421 FHSSYAHQSGLDRLNPLIDQYLYLNRTQNGSGAQNKDILFSGRSPAGMSVQPKMWLP 480
Db      421 FHSSYAHQSGLDRLNPLIDQYLYLNRTQNGSGAQNKDILFSGRSPAGMSVQPKMWLP 480
Qy      481 GPCYRQQRVSKTKTDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDDKPFPMGCV 540
Db      541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFGTVAANFOSSSTDPATGVDVHANG 600
Qy      541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFGTVAANFOSSSTDPATGVDVHANG 600
Db      601 ALPGMWQDRDYLLOCPITAKIPHTDGHFHPSPFMGREGKXPPPIILKATNPVPANRPA 660
Qy      601 ALPGMWQDRDYLLOCPITAKIPHTDGHFHPSPFMGREGKXPPPIILKATNPVPANRPA 660
Db      661 EBSATKFAFITQYSTGVSVLEIEMELQKENSGRNPEVOYTSNVAKSANVDFTVNNGI 720
Qy      721 YTEPRPIGTRYLTRPL 736
Db      721 YTEPRPIGTRYLTRPL 736

```

RESULT 11

```

ID      ADE76566
ID      ADE76566 standard; protein; 736 AA.
AC      ADE76566;
DT      29-JAN-2004 (first entry)
XX      Adeno-associated virus (AAV) related protein, SEQ ID NO 65.
XX      adeno-associated virus; AAV; cytosolic; antiposietic; antineumatic;
XX      antiairitic; neuroprotective; antidiabetic; antithyroid;
XX      dermatological; antiinflammatory; gene therapy; vaccine;
XX      hyperproliferative; cancer; psoriasis; autoimmune disease;
XX      rheumatoid arthritis; multiple sclerosis; diabetes;
XX      autoimmune thyroiditis; scleroderma; Crohn's disease.
OS      Unidentified.
XX      EP1310571-A2.
XX      14-MAY-2003.
XX      12-NOV-2002; 2002EP-00257826.
XX

```

```

PR      13-NOV-2001; 2001US-0350607P.
PR      17-DEC-2001; 2001US-0341117P.
PR      01-MAY-2002; 2002US-0377066P.
PR      05-JUN-2002; 2002US-0386675P.
PA      (UYPE-) UNIV PENNSYLVANIA.
PI      Gao G, Wilson JM, Alvira M;
DR      WPI; 2003-450984/43.
XX
PT      Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT      preventing or treating hyperproliferative or autoimmune disease,
PT      comprises subjecting a sample having a DNA to amplification via
PT      polymerase chain reaction.
XX
PS      Disclosure; SEQ ID NO 65; 419pp; English.
XX
CC      The invention relates to a novel method for detecting adeno-associated
CC      virus (AAV) sequences in a sample, which comprises subjecting a sample
CC      containing a DNA to amplification via a polymerase chain reaction (PCR).
CC      The AAV sequence have the following activities: cytosolic,
CC      antiposietic, antineumatic, antiairitic, neuroprotective,
CC      antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC      sequence can be used in gene therapy or as part of a vaccine to treat
CC      disorders. The method is useful in detecting and/or identifying AAV
CC      sequences and isolating novel sequences that are identified. The
CC      sequences may be used e.g. for preventing or treating hyperproliferative
CC      conditions such as cancers and psoriasis, and other autoimmune diseases
CC      like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC      thyroiditis, scleroderma or Crohn's disease. This sequence represents an
CC      AAV related protein sequence of the invention.
XX
SQ      Sequence 736 AA;
XX
Qy      Quezly Match 99.3%; Score 3963; DB 7; Length 736;
Qy      Best Local Similarity 99.2%; Pred. No. 6; 1e-309;
Db      Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy      1 MAADYLPDMEDNLSEGIREWMDLKPGAPKPKANQKODDGRGLVPGYKYLGPNGID 60
Db      1 MAADYLPDMEDNLSEGIREWMDLKPGAPKPKANQKODDGRGLVPGYKYLGPNGID 60
Qy      61 KGEVNAADAALEHDKAYDQOLKAGDNPYLRYNADAEFORLEEDTSFGGNLGRAVFO 120
Db      61 KGEVNAADAALEHDKAYDQOLKAGDNPYLRYNADAEFORLEEDTSFGGNLGRAVFO 120
Qy      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEOSPOBPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Db      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEOSPOBPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
Qy      181 SVPDPQPLGEPPTATPAVGPPTMASGGGAPMADNNEGADGVGNAGNHCDSITWLGDRVI 240
Db      181 SVPDPQPLGEPPTATPAVGPPTMASGGGAPMADNNEGADGVGNAGNHCDSITWLGDRVI 240
Qy      241 TTSTRTMALPTYNNHLYKQISSASTGASNDNHFGYSTWGWGFDNRFCHSPRMDORL 300
Db      241 TTSTRTMALPTYNNHLYKQISSASTGASNDNHFGYSTWGWGFDNRFCHSPRMDORL 300
Qy      301 INNMGFPKRLNFKLFNIOVKEVTTNDGVTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
Db      301 INNMGFPKRLNFKLFNIOVKEVTTNDGVTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
Qy      361 GCLPPFPADVPMIPOYGYLTLLNGSQAVERSSFYCLEYPPSQMLRTGNFTFSYTFEYV 420
Db      361 GCLPPFPADVPMIPOYGYLTLLNGSQAVERSSFYCLEYPPSQMLRTGNFTFSYTFEYV 420
Qy      421 FHSSYAHQSGLDRLNPLIDQYLYLNRTQNGSGAQNKDILFSGRSPAGMSVQPKMWLP 480
Db      421 FHSSYAHQSGLDRLNPLIDQYLYLNRTQNGSGAQNKDILFSGRSPAGMSVQPKMWLP 480
Qy      481 GPCYRQQRVSKTKTDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDDKPFPMGCV 540
Db      481 GPCYRQQRVSKTKTDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDDKPFPMGCV 540

```

Db 481 GPCYRQGVSKTKTDNNNSNFTWTGASKYKLNIGRESIIINPTAMASHKDDKFFPMGSGV 540
QY MIFGKESAGSNTALDNNMTDEBEIKATNPATERFGTVAVNPSSSTDPATGDVHAMG 600
Db 541 MIFGKESAGSNTALDNNMTDEBEIKATNPATERFGTVAVNPSSSTDPATGDVHAMG 600
QY 601 ALPGVWQDRDVLVLOGPIWAKI PHTDGHFHSPLMGFGGLKNPFPQILIKNTVPANPPA 660
Db 601 ALPGVWQDRDVLVLOGPIWAKI PHTDGHFHSPLMGFGGLKNPFPQILIKNTVPANPPA 660
QY 661 EFSATKFSFTTQVSTGVSVIEIWELOKENSCKMNPVQYTSNYSKASANDFTYDNNGL 720
Db 661 EFSATKFSFTTQVSTGVSVIEIWELOKENSCKMNPVQYTSNYSKASANDFTYDNNGL 720
QY 721 YTBPRPIGTRVLRPL 736
Db 721 YTBPRPIGTRVLRPL 736
RESULT 12
ADV70293
ID ADV70293 standard; protein; 736 AA.
XX
AC ADV70293;
XX
DT 10-MAR-2005 (first entry)
XX
DE Primate adeno-associated virus 6 capsid protein VPI.
XX
KW immunosuppressive; gene therapy; immunogenicity; virus inactivation;
KW hemophilia; peptic deficiency; galactosemia; phenylketonuria;
KW Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
KW atherosclerosis; thrombosis; embolism; Parkinson's disease;
KW congestive heart failure; cancer; inflammation; immune disorder;
KW muscular dystrophy; diabetes; VPI.
XX
OS Adeno-associated virus 6.
XX
PN MO2004112727-A2.
XX
PD 29-DEC-2004.
XX
PF 21-JUN-2004; 2004MO-US019884.
XX
PR 19-JUN-2003; 2003US-0480395P.
PR 30-APR-2004; 2004US-0567310P.
PR 03-JUN-2004; 2004US-0576501P.
XX
PA (AVIG-) AVIGEN INC.
XX
PI Arbetman AE, Colosi P, Lochrie MA, Surosky RT;
XX
DR WPI, 2005-048755/05.
XX
PT New mutated adeno-associated virus (AAV) capsid protein that when present
PT in an AAV virion imparts decreased immunoreactivity to the virion as
PT compared to the corresponding wild-type virion, useful for treating e.g.
PT hemophilia.
XX
PS Example 5; SEQ ID NO 19; 136pp; English.
XX
CC The invention describes a mutated adeno-associated virus (AAV) capsid
CC protein that when present in an AAV virion imparts decreased
CC immunoreactivity to the virion as compared to the corresponding wild-type
CC virion. Also described are: a polynucleotide encoding the mutated protein
CC above; a recombinant AAV virion comprising the mutated protein above; and
CC delivering a recombinant AAV virion to a cell or tissue of a vertebrate
CC subject. The recombinant AAV virion is useful for delivering a
CC heterologous nucleic acid molecule to a cell or tissue of a vertebrate
CC subject, where the protein encoded by the heterologous nucleic acid
CC molecule is expressed at a level that provides a therapeutic effect.
CC where the recombinant AAV virion may comprise a non-primate, mammalian
CC AAV capsid protein that when present in an AAV virion imparts decreased

CC immunoreactivity to the virion as compared to immunoreactivity of primate
CC AAV-2, and the heterologous nucleic acid molecule, where the heterologous
CC nucleic acid molecule encodes a therapeutic protein and is operably
CC linked to control elements capable of directing the in vivo transcription
CC and translation of the protein. The protein or the recombinant AAV virion
CC is useful for treating or preventing a wide variety of disorders such as
CC hemophilia, glycogen storage deficiency type 1A, peptic deficiency,
CC galactosemia, phenylketonuria, Crigler-Najjar disease, Gout and Lesch-
CC Nyman syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,
CC thrombosis, embolisms, Parkinson's disease, congestive heart failure,
CC cancer, inflammatory and immune disorders, muscular dystrophies, and
CC diabetes. This is the amino acid sequence of adeno-associated virus 6
CC (AAV6) capsid protein VPI.
XX
SQ Sequence 736 AA;

Query Match 99.3%; Score 3963; DB 9; Length 736;
Best Local Similarity 99.2%; Pred. No. 6.1e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAADGYLPMWLENDLSEGIREFMMDLKPQAPKPKANQKODDGRGLVPGYKYLGPFGNGL 60
Db 1 MAADGYLPMWLENDLSEGIREFMMDLKPQAPKPKANQKODDGRGLVPGYKYLGPFGNGL 60
QY 61 KGEFVNADAAALHEDKAYDQOLKAGDNPYLRYNHADAEFGERLQEDTSPFGNIGRAVFQ 120
Db 61 KGEFVNADAAALHEDKAYDQOLKAGDNPYLRYNHADAEFGERLQEDTSPFGNIGRAVFQ 120
QY 121 AKKRVLEPLGLVERGAKTAPKKRPVQSPQEPSSSGIGRTGQOPAKKRLNFCQTDSE 180
Db 121 AKKRVLEPLGLVERGAKTAPKKRPVQSPQEPSSSGIGRTGQOPAKKRLNFCQTDSE 180
QY 181 SVDPQPLGEPFPAVAVGPTTMSGGGAPADNNEGADGVGNSGWNHCDSTWLGDRVI 240
Db 181 SVDPQPLGEPFPAVAVGPTTMSGGGAPADNNEGADGVGNSGWNHCDSTWLGDRVI 240
QY 241 TTSRTMALPTYNHHLVKQISSASTGASNDNHVFGYSTPMWGFDFNRFHCFSPRDWQL 300
Db 241 TTSRTMALPTYNHHLVKQISSASTGASNDNHVFGYSTPMWGFDFNRFHCFSPRDWQL 300
QY 301 INNNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVPSDSRYQLPYVLGSAHQ 360
Db 301 INNNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVPSDSRYQLPYVLGSAHQ 360
QY 361 GCLPPFPAVDFMIQOYGLTLNNGSQAVGRSPFCLETFPSQMLRTGNPFTFTFEEDVP 420
Db 361 GCLPPFPAVDFMIQOYGLTLNNGSQAVGRSPFCLETFPSQMLRTGNPFTFTFEEDVP 420
QY 421 FHSSVYAHQSILDRMLNPLIDQYLYLNRTONQSGSAQNKDILFSRGSFAGMSVOPKMWLP 480
Db 421 FHSSVYAHQSILDRMLNPLIDQYLYLNRTONQSGSAQNKDILFSRGSFAGMSVOPKMWLP 480
QY 481 GPCYRQGVSKTKTDNNNSNFTWTGASKYKLNIGRESIIINPTAMASHKDDKFFPMGSGV 540
Db 481 GPCYRQGVSKTKTDNNNSNFTWTGASKYKLNIGRESIIINPTAMASHKDDKFFPMGSGV 540
QY 541 MIFGKESAGSNTALDNNMTDEBEIKATNPATERFGTVAVNPSSSTDPATGDVHAMG 600
Db 541 MIFGKESAGSNTALDNNMTDEBEIKATNPATERFGTVAVNPSSSTDPATGDVHAMG 600
QY 601 ALPGVWQDRDVLVLOGPIWAKI PHTDGHFHSPLMGFGGLKNPFPQILIKNTVPANPPA 660
Db 601 ALPGVWQDRDVLVLOGPIWAKI PHTDGHFHSPLMGFGGLKNPFPQILIKNTVPANPPA 660
QY 661 EFSATKFSFTTQVSTGVSVIEIWELOKENSCKMNPVQYTSNYSKASANDFTYDNNGL 720
Db 661 EFSATKFSFTTQVSTGVSVIEIWELOKENSCKMNPVQYTSNYSKASANDFTYDNNGL 720
QY 721 YTBPRPIGTRVLRPL 736
Db 721 YTBPRPIGTRVLRPL 736

RESULT 13

AD227070
ID AD227070 standard; protein; 736 AA.

AC AD227070;

XX 30-JUN-2005 (first entry)

DE Adeno-associated virus protein SEQ ID NO 220.

XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX bacterial infection; cancer; ulcerative colitis; antineoplastic;
XX antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
XX antipneumonia; vasotropic; gastrointestinal; gen.; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antitumor; dermatological.

OS Adeno-associated virus.

XX WO2005033321-A2.

XX 14-APR-2005.

XX 30-SEP-2004; 2004WO-US028817.

XX 30-SEP-2003; 2003US-0508226P.

XX 29-APR-2004; 2004US-0566546P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX WPI; 2005-285437/29.

XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.

XX Disclosure; SEQ ID NO 220; 569pp; English.

XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents the amino acid sequence of an
XX adeno-associated virus protein.

XX Sequence 736 AA;

XX Query Match 99.3%; Score 3963; DB 9; Length 736;

XX Best Local Similarity 99.2%; Pred. No. 6.1e-309;
XX Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

XX 1 MAADGYLPDWLENDLSEGIKREWMDLKPGAPKPRANQKQDDGRLVPGYKYLGPFGD 60

XX 1 MAADGYLPDWLENDLSEGIKREWMDLKPGAPKPRANQKQDDGRLVPGYKYLGPFGD 60

XX 61 KGEFVNAADAAAEHDKAYVOOLKAGNPLYRNHDAEFOERLQEDTSFGNLSGRAVQ 120

XX 61 KGEFVNAADAAAEHDKAYVOOLKAGNPLYRNHDAEFOERLQEDTSFGNLSGRAVQ 120

XX 121 AKKRVLEPLGIVEGAKTAPKRRPVESPOEPDSSSGIKTGQPAKKRLNFGQTGDS 180

XX 121 AKKRVLEPLGIVEGAKTAPKRRPVESPOEPDSSSGIKTGQPAKKRLNFGQTGDS 180

XX 181 SVDPDPLGEPDPATPAVGPPTMASGGAPMADNNEGADGVGNAGNHCDSITWLGDRV 240

DB 181 SVDPDPLGEPDPATPAVGPPTMASGGAPMADNNEGADGVGNAGNHCDSITWLGDRV 240

XX 241 TTSITRTMALPTYYNNHLYQOISSASTGASNDNHYFGYSTPMGTFEDNRHCHSPDDKORL 300

XX 241 TTSITRTMALPTYYNNHLYQOISSASTGASNDNHYFGYSTPMGTFEDNRHCHSPDDKORL 300

XX 301 INNMGFPKRLNPLFNLQYKVEVTTNDGVTIANNLTSTQVFSDESXQLPYVLSGAHQ 360

XX 301 INNMGFPKRLNPLFNLQYKVEVTTNDGVTIANNLTSTQVFSDESXQLPYVLSGAHQ 360

XX 361 GCLPFPADVFMIPQYGLTLTNGSQAVGRSFLCYLFPQMLRTGNNFTFSYFEEVP 420

XX 361 GCLPFPADVFMIPQYGLTLTNGSQAVGRSFLCYLFPQMLRTGNNFTFSYFEEVP 420

XX 421 FHSSTAHQSGLDRLNPLIDQYLYLNTQONOSGAONKDLLFSRSGPAGMSVQKNTLP 480

XX 421 FHSSTAHQSGLDRLNPLIDQYLYLNTQONOSGAONKDLLFSRSGPAGMSVQKNTLP 480

XX 421 FHSSTAHQSGLDRLNPLIDQYLYLNTQONOSGAONKDLLFSRSGPAGMSVQKNTLP 480

XX 481 GPCYRQORVSKTKTNNNSNFTWTGASKYKYLNGRESIINPTAMASHKODEKFPMSGV 540

XX 481 GPCYRQORVSKTKTNNNSNFTWTGASKYKYLNGRESIINPTAMASHKODEKFPMSGV 540

XX 541 MIFGESAGASNTALDNMTDDEEIKATNPVATERFGTVAVNFOSSSTDPATGCVHANG 600

XX 541 MIFGESAGASNTALDNMTDDEEIKATNPVATERFGTVAVNFOSSSTDPATGCVHANG 600

XX 601 ALPGWVQDRDYLQGPIMAKIPHTDGHFHPPLMGFGGLKPPQIILIKNTVPVNPAP 660

XX 601 ALPGWVQDRDYLQGPIMAKIPHTDGHFHPPLMGFGGLKPPQIILIKNTVPVNPAP 660

XX 661 EFSATKFPASFTIOYSTGVSEIEMWELQENSKRNPEVQYTSNYSKANSANVPTVNNGL 720

XX 661 EFSATKFPASFTIOYSTGVSEIEMWELQENSKRNPEVQYTSNYSKANSANVPTVNNGL 720

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

XX 721 YTEPRPIGRTYLRPL 736

PT New adeno-associated virus (AAV) clade comprising at least three AAV
 CC members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.

XX Claim 19; SEQ ID NO 236; 569pp; English.

CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.

CC Sequence 737 AA;

Query Match 97.7%; Score 3897.5; DB 9; Length 737;

Best Local Similarity 97.7%; Pred. No. 1.1e-303; Indels 1; Gaps 1;
 Matches 720; Conservative 6; Mismatches 10;

```

OY 1 MAADGYLPMWLEDNUSGIRREWMDLKPAPKPKANQKODDGRGLVLPGYKYLGPFGNLD 60
DB 1 MAADGYLPMWLEDNUSGIRREWMDLKPAPKPKANQKODDGRGLVLPGYKYLGPFGNLD 60
OY 61 KGEPVNAADAALBHDKAYDQQLKAGDNPYLRYNHADAEPORLOEDTSFGNIGRAVFO 120
DB 61 KGEPVNAADAALBHDKAYDQQLKAGDNPYLRYNHADAEPORLOEDTSFGNIGRAVFO 120
OY 121 AKKRVLEPLGLVEBKAATAPGKKRPVEQSPQ-EPDSSSGIGCTGQOPAKKLNFGQTGDS 179
DB 121 AKKRVLEPLGLVEBKAATAPGKKRPVEQSPQ-EPDSSSGIGCTGQOPAKKLNFGQTGDS 180
OY 180 ESVPPDPLGEPBPAPPAVPTTMAAGGAPMADNNEGADGVGNAGMWHCDSTLGRV 239
DB 181 ESVPPDPLGEPBPAPPAVPTTMAAGGAPMADNNEGADGVGNAGMWHCDSTLGRV 240
OY 240 ITTSTRMALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYPFDFNPFHCFSPRDMOR 299
DB 241 ITTSTRMALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYPFDFNPFHCFSPRDMOR 300
OY 300 LINNMWGRPRKLNKLPRIOYKEVTTNDGVTIANNLSTVOVPSDBEYOLPYVLGSAH 359
DB 301 LINNMWGRPRKLNKLPRIOYKEVTTNDGVTIANNLSTVOVPSDBEYOLPYVLGSAH 360
OY 360 OGCLPPFPADVPMIYOYGYLTNNGSOAVGRSFCLETFPSPQMLRTGNNFTFSYTFEBV 419
DB 361 OGCLPPFPADVPMIYOYGYLTNNGSOAVGRSFCLETFPSPQMLRTGNNFTFSYTFEBV 420
OY 420 PFHSSVYAHQSJDLRLMNPIDQYLYLNRTONQSSAQNKDLLFSGSPAGMSVOPKMWL 479
DB 421 PFHSSVYAHQSJDLRLMNPIDQYLYLNRTONQSSAQNKDLLFSGSPAGMSVOPKMWL 480
OY 480 BGPCTVROORVSTKTDNNNSNFTWVGASKYNLNGEESIINPTAASHDDDKFPFMSG 539
DB 481 BGPCTVROORVSTKTDNNNSNFTWVGASKYNLNGEESIINPTAASHDDDKFPFMSG 540
OY 540 VMIPEGESAGSNTALDNVMTDEBEIKATNPVATERFCTVAVNFOSSSTDATGDVAM 599
DB 541 VMIPEGESAGSNTALDNVMTDEBEIKATNPVATERFCTVAVNFOSSSTDATGDVAM 600
OY 600 GALPGWMDRDUVYLGPIWAKIPHTDGHFSPMLMGFGKLNPPQILIKNTPVANPP 659
DB 601 GALPGWMDRDUVYLGPIWAKIPHTDGHFSPMLMGFGKLNPPQILIKNTPVANPP 660
OY 660 AEFSAATKFASTITTOYSTGOVSVEIEMELOKENSKWNPVOTSTYASANDFTVDNNG 719
DB 661 AEFSAATKFASTITTOYSTGOVSVEIEMELOKENSKWNPVOTSTYASANDFTVDNNG 720

```

OY 720 LYTEPRPIGTRYLTRPL 736
 DB 721 LYTEPRPIGTRYLTRPL 737

RESULT 15

ID AD227010 standard; protein; 737 AA.

AC AD227010;

DT 30-JUN-2005 (first entry)

DE Adeno-associated virus protein SEQ ID NO 160.

XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
 KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
 KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 KW virucide; antibacterial; cytostatic; antiulcer; dermatological.

OS Adeno-associated virus.

PN WO200503321-A2.

PD 14-APR-2005.

PF 30-SEP-2004; 2004WO-US028817.

PR 30-SEP-2003; 2003US-0508226P.

PR 29-APR-2004; 2004US-0566546P.

PA (UYPE-) UNIV PENNSYLVANIA.

PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

DR WPI; 2005-285437/29.

PT New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.

PS Claim 46; SEQ ID NO 160; 569pp; English.

CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.

XX Sequence 737 AA;

Query Match 97.4%; Score 3883.5; DB 9; Length 737;

Best Local Similarity 97.4%; Pred. No. 1.5e-302; Indels 1; Gaps 1;
 Matches 718; Conservative 6; Mismatches 12;

```

OY 1 MAADGYLPMWLEDNUSGIRREWMDLKPAPKPKANQKODDGRGLVLPGYKYLGPFGNLD 60
DB 1 MAADGYLPMWLEDNUSGIRREWMDLKPAPKPKANQKODDGRGLVLPGYKYLGPFGNLD 60
OY 61 KGEPVNAADAALBHDKAYDQQLKAGDNPYLRYNHADAEPORLOEDTSFGNIGRAVFO 120
DB 61 KGEPVNAADAALBHDKAYDQQLKAGDNPYLRYNHADAEPORLOEDTSFGNIGRAVFO 120
OY 121 AKKRVLEPLGLVEBKAATAPGKKRPVEQSPQ-EPDSSSGIGCTGQOPAKKLNFGQTGDS 179

```

```

Db      ||||| 121 AKKRVLEPLGLVEAKTAPEGKRPEPSPORS PDSSTIGKKGOQPAKKRLNFGQTGDS 180
Qy      ||||| 180 ESVDPDQPIGEPBPATPAVAVPTTMAAGGAPMADNNEGADGVNAGSNWHCDSTWLGDRV 239
Db      ||||| 181 ESVDPDQPIGEPBPAGSGISGTMAGGAPMADNNEGADGVNAGSNWHCDSTWLGDRV 240
Qy      ||||| 240 ITTSTRTMALPTYNNHLYKQISSASTGASNDNHFGYSTPWGYPFDFNRPHCHFSPRDMOR 299
Db      ||||| 241 ITTSTRTMALPTYNNHLYKQISSASTGASNDNHFGYSTPWGYPFDFNRPHCHFSPRDMOR 300
Qy      ||||| 300 LNNMGPRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTVQVFSDEYOLPYVLGSAH 359
Db      ||||| 301 LNNMGPRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTVQVFSDEYOLPYVLGSAH 360
Qy      ||||| 360 OGCLPFPADVFMIPOYGYLTLLNGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEV 419
Db      ||||| 361 OGCLPFPADVFMIPOYGYLTLLNGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEV 420
Qy      ||||| 420 PFHSSYAHSQSLDRLMNPLIDQYLYYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKMWL 479
Db      ||||| 421 PLHSSYAHSQSLDRLMNPLIVQYLYYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKMWL 480
Qy      ||||| 480 PGFCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 539
Db      ||||| 481 PGFCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
Qy      ||||| 540 VMI FGKESAGASNTALDNVMI TDEEBEIKATNPVATERFGTVAVNFOSSSTDPAIGDVHAM 599
Db      ||||| 541 VMI FGKESAGASNTALDNVMI TDEEBEIKATNPVATERFGTVAVNFOSSSTDPAIGDVHAM 600
Qy      ||||| 600 GALPGWVWODRDVYLOGPIWAKI PHTDGHFHPSPLMGGFGLKNP PPQILIKNTVPANPP 659
Db      ||||| 601 GALPGWVWODRDVYLOGPIWAKI PHTDGHFHPSPLMGGFGLKNP PPQILIKNTVPANPP 660
Qy      ||||| 660 AEFSAITKFASTIQTSTGVSVIELEWLOKENSKRNNPEVQYTSNAXSANDVFTVDNNG 719
Db      ||||| 661 AEFSAITKFASTIQTSTGVSVIELEWLOKENSKRNNPEVQYTSNAXSANDVFTVDNNG 720
Qy      ||||| 720 LYTEPRPIGTRYLTRPL 736
Db      ||||| 721 LYTEPRPIGTRYLTRPL 737

```

Search completed: November 23, 2005, 17:35:10
 Job time : 134.527 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compen Ltd.

OM protein - protein search, using SW model

Run on: November 23, 2005, 17:26:29 ; Search time 33.4724 Seconds
(without alignments)
1817.892 Million cell updates/sec

Title: US-10-696-282-13

Perfect score: 3989

Sequence: 1 MAADGTLPMLEDNLSGIR.....NNGLYTEPRPIGTRVLTPL 736

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents_AA:*

- 1: /cgn2_6/prodata/1/iaa/5.COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/6.COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/H.COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/RE.COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfillset1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3989	100.0	736	2	US-09-807-802A-3
2	3989	100.0	736	2	US-09-807-802A-13
3	3402.5	85.3	735	2	US-09-321-589-1
4	3402.5	85.3	735	2	US-10-293-478-1
5	3402.5	85.3	735	2	US-10-038-972A-13
6	3251	81.5	734	2	US-09-807-802A-17
7	2906	72.9	534	2	US-09-807-802A-17
8	2759.5	69.2	538	2	US-10-038-972A-14
9	2486.5	62.2	734	2	US-09-532-594B-4
10	2481.5	62.2	533	2	US-10-038-972A-15
11	2220	55.7	724	2	US-09-533-427-4
12	1830.5	45.9	598	2	US-09-532-594B-16
13	1700.5	42.6	588	2	US-09-533-427-5
14	1690.5	42.4	544	2	US-09-532-594B-18
15	1665	41.7	532	2	US-09-533-427-6
16	1410	35.3	756	2	US-09-438-268-4
17	599.5	15.0	781	2	US-10-187-253E-27
18	590.5	14.8	781	2	US-10-187-253E-27
19	486	12.2	554	2	US-10-187-253E-29
20	486	12.2	554	2	US-10-187-253E-33
21	479.5	12.0	543	2	US-08-856-841-22
22	430.5	10.8	500	2	US-08-856-841-16
23	430.5	10.8	501	2	US-08-856-841-19
24	429.5	10.8	486	2	US-08-856-841-19
25	326	8.2	415	2	US-08-856-841-20
26	318	8.0	395	2	US-08-856-841-13
27	310	7.8	264	2	US-08-856-841-14

28	303	7.6	398	2	US-08-856-841-21	Sequence 21, Appl
29	284	7.1	387	2	US-08-856-841-17	Sequence 17, Appl
30	216.5	5.4	579	6	5223424-13	Patent No. 5223424
31	192	4.8	584	2	US-09-022-949-2	Sequence 2, Appl
32	142	3.6	210	2	US-08-856-841-9	Sequence 9, Appl
33	142	3.6	227	2	US-08-856-841-15	Sequence 15, Appl
34	142	3.6	250	2	US-08-856-841-12	Sequence 12, Appl
35	132.5	3.3	489	2	US-10-376-397B-4	Sequence 4, Appl
36	124	3.1	3060	1	US-08-487-826B-14	Sequence 14, Appl
37	122.5	3.1	1637	2	US-10-172-502-14	Sequence 14, Appl
38	121	3.0	2736	2	US-09-252-991A-30227	Sequence 30227, A
39	119.5	3.0	1394	2	US-08-296-791-2	Sequence 2, Appl
40	119.5	3.0	1394	2	US-09-839-996-2	Sequence 2, Appl
41	119.5	3.0	1394	2	US-10-080-505-2	Sequence 2, Appl
42	119.5	3.0	1394	2	US-10-645-655-2	Sequence 2, Appl
43	119.5	3.0	1394	4	PCT-US95-10661A-2	Sequence 2, Appl
44	119	3.0	1095	2	US-09-107-532A-3855	Sequence 3855, Ap
45	117.5	2.9	1395	2	US-10-080-505-7	Sequence 7, Appl

ALIGNMENTS

```
RESULT 1
US-09-807-802A-3
; Sequence 3, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Xiao, Weidong
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: GNVN.03105A
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-3

Query Match      100.0%; Score 3989; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAADGTLPMLEDNLSGIRREWMWLKPKAKNOQKDDGRLVPGYKYLGPNGLD 60
        |||
Db      1 MAADGTLPMLEDNLSGIRREWMWLKPKAKNOQKDDGRLVPGYKYLGPNGLD 60

QY      1 MAADGTLPMLEDNLSGIRREWMWLKPKAKNOQKDDGRLVPGYKYLGPNGLD 60
        |||
Db      1 MAADGTLPMLEDNLSGIRREWMWLKPKAKNOQKDDGRLVPGYKYLGPNGLD 60

QY      61 KGEPVNAADAAIEHDKAYDQOLKAGDNPLYRNNHDAEFOEHLQDTSFGNVLGAHVQ 120
        |||
Db      61 KGEPVNAADAAIEHDKAYDQOLKAGDNPLYRNNHDAEFOEHLQDTSFGNVLGAHVQ 120

QY      121 AKRRVLEPIGLVEGAKTAPGKKRPVEQSPQEDSSSGIGKTGQOPAKKRLNFGQTDSE 180
        |||
Db      121 AKRRVLEPIGLVEGAKTAPGKKRPVEQSPQEDSSSGIGKTGQOPAKKRLNFGQTDSE 180

QY      181 SVDPQPLGEPATPAVGPFTTMAASGCGAPMADNNEGADGVGASGNMHCDSITWLGDRVT 240
        |||
Db      181 SVDPQPLGEPATPAVGPFTTMAASGCGAPMADNNEGADGVGASGNMHCDSITWLGDRVT 240

QY      241 TTSTRFALPTVNNHLVKQISSASTASNDNHFGYSTPWGTFDNRFFCHSPRQWL 300
        |||
Db      241 TTSTRFALPTVNNHLVKQISSASTASNDNHFGYSTPWGTFDNRFFCHSPRQWL 300

QY      301 INNNKFRPRRLNFKLFTNVQKVTNNDDVTTANNLTSTGVQFSDSEYQLPVVLASAHQ 360
        |||
Db      301 INNNKFRPRRLNFKLFTNVQKVTNNDDVTTANNLTSTGVQFSDSEYQLPVVLASAHQ 360
```

```
Db 301 INNMGPRPKRLNFKLFINIYQVKEVTTNDGVTIANLSTVQVFSDEYQLPYVLGSAHQ 360
Qy 361 GCLPPFPADVMTIPQYGYLTLLNNGSOAVGRSSFCLEYPFSQMLRTGNFFSTFESEVP 420
Db 361 GCLPPFPADVMTIPQYGYLTLLNNGSOAVGRSSFCLEYPFSQMLRTGNFFSTFESEVP 420
Qy 421 FHSYASQSOLDRMLNPLIDQYLYLNRTONOSGAONKDLFBSRGPAGMSVOPKMWLP 480
Db 421 FHSYASQSOLDRMLNPLIDQYLYLNRTONOSGAONKDLFBSRGPAGMSVOPKMWLP 480
Qy 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGV 540
Db 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGV 540
Qy 541 MIFGKESAGASNTALDNNMTIDESEIKATNPVATERGTYAVNFOSSSTDPATGDVHAMG 600
Db 541 MIFGKESAGASNTALDNNMTIDESEIKATNPVATERGTYAVNFOSSSTDPATGDVHAMG 600
Qy 601 ALPGWVODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTPVANPAPA 660
Db 601 ALPGWVODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTPVANPAPA 660
Qy 661 EFSATKFSFITYSTGQVSVIEIWELOKENSKRKNPEVOYTSNYAKSANDFTVDNNGL 720
Db 661 EFSATKFSFITYSTGQVSVIEIWELOKENSKRKNPEVOYTSNYAKSANDFTVDNNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736
```

RESULT 2

```
US-09-807-802A-13
; Sequence 13, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiac, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.0310USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/407,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-13
```

```
Query Match 100.0%; Score 3989; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MAADGYLPDWLEDTLSEGIREWMDLKPQAPKPKANQKQDDGRGLVLPQYKYLGFPGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIREWMDLKPQAPKPKANQKQDDGRGLVLPQYKYLGFPGNGLD 60
Qy 61 KEEPNNAADAALBHDKAYDQOLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFO 120
Db 61 KEEPNNAADAALBHDKAYDQOLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFO 120
Qy 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGQOPAKKRLNFGQTGSE 180
Db 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGQOPAKKRLNFGQTGSE 180
Qy 181 SVDPDQPLGEPATPAVGPITMASGGGAPMADNNEGADGVNAGNMHCDSTWLGDYVI 240
Db 181 SVDPDQPLGEPATPAVGPITMASGGGAPMADNNEGADGVNAGNMHCDSTWLGDYVI 240
```

```
Db 181 SVDPDQPLGEPATPAVGPITMASGGGAPMADNNEGADGVNAGNMHCDSTWLGDYVI 240
Qy 241 TTSRTALPPTYNNHVKQISSASTGASNDNHVGYSTPMGYDPFNRFHCFSPRDMORL 300
Db 241 TTSRTALPPTYNNHVKQISSASTGASNDNHVGYSTPMGYDPFNRFHCFSPRDMORL 300
Qy 301 INNMGPRPKRLNFKLFINIYQVKEVTTNDGVTIANLSTVQVFSDEYQLPYVLGSAHQ 360
Db 301 INNMGPRPKRLNFKLFINIYQVKEVTTNDGVTIANLSTVQVFSDEYQLPYVLGSAHQ 360
Qy 361 GCLPPFPADVMTIPQYGYLTLLNNGSOAVGRSSFCLEYPFSQMLRTGNFFSTFESEVP 420
Db 361 GCLPPFPADVMTIPQYGYLTLLNNGSOAVGRSSFCLEYPFSQMLRTGNFFSTFESEVP 420
Qy 421 FHSYASQSOLDRMLNPLIDQYLYLNRTONOSGAONKDLFBSRGPAGMSVOPKMWLP 480
Db 421 FHSYASQSOLDRMLNPLIDQYLYLNRTONOSGAONKDLFBSRGPAGMSVOPKMWLP 480
Qy 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGV 540
Db 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGV 540
Qy 541 MIFGKESAGASNTALDNNMTIDESEIKATNPVATERGTYAVNFOSSSTDPATGDVHAMG 600
Db 541 MIFGKESAGASNTALDNNMTIDESEIKATNPVATERGTYAVNFOSSSTDPATGDVHAMG 600
Qy 601 ALPGWVODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTPVANPAPA 660
Db 601 ALPGWVODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTPVANPAPA 660
Qy 661 EFSATKFSFITYSTGQVSVIEIWELOKENSKRKNPEVOYTSNYAKSANDFTVDNNGL 720
Db 661 EFSATKFSFITYSTGQVSVIEIWELOKENSKRKNPEVOYTSNYAKSANDFTVDNNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736
```

RESULT 3

```
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1
```

```
Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.2e-250;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
```

```
Qy 1 MAADGYLPDWLEDTLSEGIREWMDLKPQAPKPKANQKQDDGRGLVLPQYKYLGFPGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIREWMDLKPQAPKPKANQKQDDGRGLVLPQYKYLGFPGNGLD 60
Qy 61 KEEPNNAADAALBHDKAYDQOLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFO 120
Db 61 KEEPNNAADAALBHDKAYDQOLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFO 120
Qy 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGQOPAKKRLNFGQTGSE 180
Db 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGQOPAKKRLNFGQTGSE 180
Qy 181 SVDPDQPLGEPATPAVGPITMASGGGAPMADNNEGADGVNAGNMHCDSTWLGDYVI 240
Db 181 SVDPDQPLGEPATPAVGPITMASGGGAPMADNNEGADGVNAGNMHCDSTWLGDYVI 240
```



```
QY 181 SVDPQPLGEPATPAVGPPTTMASGGGAPADNNEGADVGNASGNHCDSTWLGDRVI 240
DB 181 SVDPQPLGEPATPAVGPPTTMASGGGAPADNNEGADVGNASGNHCDSTWLGDRVI 240
QY 241 TTSTRTMALPTYNHNLHYKQISSASTGASNDNHFGYSTPMGYFDFNRPHCHSPRDWQRL 300
DB 241 TTSTRTMALPTYNHNLHYKQISSOS - GASNDNHFGYSTPMGYFDFNRPHCHSPRDWQRL 299
QY 301 INNMGFRPKRLNFKLFNIQVKEVTNDGTTIANNLSTVOVFPDSEYQLPYVLGSAHQ 360
DB 300 INNMGFRPKRLNFKLFNIQVKEVTNDGTTIANNLSTVOVFPDSEYQLPYVLGSAHQ 359
QY 361 GCLPPPADVPMIPOYGYLTLNNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEEVP 420
DB 360 GCLPPPADVPMIPOYGYLTLNNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEEVP 419
QY 421 FHSSYAHQSGLDRMLNPLIDQYLYLNRTONOSGSAONKDLLFSRGSFAGMSVQPKMWLP 480
DB 420 FHSSYAHQSGLDRMLNPLIDQYLYLNRTPSGTTQSRLOFSQAGASDIRDOSRWMLP 479
QY 481 GPCYRQORVSKTNDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
DB 480 GPCYRQORVSKTSADNNSEYSWTGATKTHLNGRDSLVPNGPAMASHKDEDEKFFPMSCV 539
QY 541 MIFGKESAGASTALDNNMTDEEBEIKATNPVATERFGTVAVNFQSSSTDPAVDYHANG 600
DB 540 LIFGKGSEKTNVDIEKWTITDEEBEIRTNVPATEQYGSVTLQGRNOAATADVNTQG 599
QY 601 ALPGMWODRDVYLGQPIWAKIPHTDGHFHPSPLMGFGFLKNPPOILLKNTPVANPAPA 660
DB 600 VLPGMWODRDVYLGQPIWAKIPHTDGHFHPSPLMGFGFLKNPPOILLKNTPVANPAPST 659
QY 661 EFSATKPAFTIOYSTGOVSVEIEMELOKENSKRMPPEVQYTSNKAASANDVFTDNNGL 720
DB 660 TFSAAKFASFITIOYSTGOVSVEIEMELOKENSKRMPPEIOTYSNKAASVANDVFTDNTGV 719
QY 721 YTEPRPIGTRYLTRPL 736
DB 720 YSEPRPIGTRYLTRNL 735

RESULT 4
US-10-293-478-1
; Sequence 1, Application US/10293478
; Patent No. 6733757
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PR1
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.2e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
```

```
QY 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQKODDGRGLVLPGYKYLGPFNGLD 60
DB 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQKODDGRGLVLPGYKYLGPFNGLD 60
QY 61 KGEVFNADAAALEHDKAYRQDLSDGNPVLKYNHADAFOERLKEDTSPFGNLGRAVFO 120
DB 61 KGEVFNADAAALEHDKAYRQDLSDGNPVLKYNHADAFOERLKEDTSPFGNLGRAVFO 120
```

```
DB 61 KGEVFNADAAALEHDKAYRQDLSDGNPVLKYNHADAFOERLKEDTSPFGNLGRAVFO 120
QY 121 AKKRLLEPLGLVEBEAKTAPKPKRPVEOSPOEPPSSGIGTKGOOPAKKRLNFGOTGSE 180
DB 121 AKKRLLEPLGLVEBEVKTAPKPKRPVEHSPVEPSSSGTGAQOQAPKRLNFGOTGAD 180
QY 181 SVDPQPLGEPATPAVGPPTTMASGGGAPADNNEGADVGNASGNHCDSTWLGDRVI 240
DB 181 SVDPQPLGEPATPAVGPPTTMASGGGAPADNNEGADVGNASGNHCDSTWLGDRVI 240
QY 241 TTSTRTMALPTYNHNLHYKQISSASTGASNDNHFGYSTPMGYFDFNRPHCHSPRDWQRL 300
DB 241 TTSTRTMALPTYNHNLHYKQISSOS - GASNDNHFGYSTPMGYFDFNRPHCHSPRDWQRL 299
QY 301 INNMGFRPKRLNFKLFNIQVKEVTNDGTTIANNLSTVOVFPDSEYQLPYVLGSAHQ 360
DB 300 INNMGFRPKRLNFKLFNIQVKEVTNDGTTIANNLSTVOVFPDSEYQLPYVLGSAHQ 359
QY 361 GCLPPPADVPMIPOYGYLTLNNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEEVP 420
DB 360 GCLPPPADVPMIPOYGYLTLNNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEEVP 419
QY 421 FHSSYAHQSGLDRMLNPLIDQYLYLNRTONOSGSAONKDLLFSRGSFAGMSVQPKMWLP 480
DB 420 FHSSYAHQSGLDRMLNPLIDQYLYLNRTPSGTTQSRLOFSQAGASDIRDOSRWMLP 479
QY 481 GPCYRQORVSKTNDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
DB 480 GPCYRQORVSKTSADNNSEYSWTGATKTHLNGRDSLVPNGPAMASHKDEDEKFFPMSCV 539
QY 541 MIFGKESAGASTALDNNMTDEEBEIKATNPVATERFGTVAVNFQSSSTDPAVDYHANG 600
DB 540 LIFGKGSEKTNVDIEKWTITDEEBEIRTNVPATEQYGSVTLQGRNOAATADVNTQG 599
QY 601 ALPGMWODRDVYLGQPIWAKIPHTDGHFHPSPLMGFGFLKNPPOILLKNTPVANPAPA 660
DB 600 VLPGMWODRDVYLGQPIWAKIPHTDGHFHPSPLMGFGFLKNPPOILLKNTPVANPAPST 659
QY 661 EFSATKPAFTIOYSTGOVSVEIEMELOKENSKRMPPEVQYTSNKAASANDVFTDNNGL 720
DB 660 TFSAAKFASFITIOYSTGOVSVEIEMELOKENSKRMPPEIOTYSNKAASVANDVFTDNTGV 719
QY 721 YTEPRPIGTRYLTRPL 736
DB 720 YSEPRPIGTRYLTRNL 735

RESULT 5
US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 2835/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PR1
; ORGANISM: adeno-associated virus 2 VP1 capsid protien
US-10-038-972A-13

Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.2e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
```

```
QY 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQKODDGRGLVLPGYKYLGPFNGLD 60
DB 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQKODDGRGLVLPGYKYLGPFNGLD 60
QY 61 KGEVFNADAAALEHDKAYRQDLSDGNPVLKYNHADAFOERLKEDTSPFGNLGRAVFO 120
DB 61 KGEVFNADAAALEHDKAYRQDLSDGNPVLKYNHADAFOERLKEDTSPFGNLGRAVFO 120
```

```
Db      1 MAADGYLPEWLEDTLSEGIROWMLKGPBPPEKPAERHKDSDRGVLPGYKYLGPFGNGLD 60
Qy      61 KGEPNNAADAALHDKAYDQDLKAGDNPYLRYNNAADAEPOERLOEDPSFGANIGRAVQ 120
Db      61 KGEPNNAADAALHDKAYDROLDSDNRPYLKYNNAADAEPOERLEEDPSFGANIGRAVQ 120
Qy      121 AKKRVLBLEGLVEBAGAKTAPGKKRPVEBOSPOEPDSSSIGIKTGOQPAKKRLNFGQTDSE 180
Db      121 AKKRVLBLEGLVEBAGAKTAPGKKRPVEBOSPOEPDSSSIGIKTGOQPAKKRLNFGQTDSE 180
Qy      181 SVDPQPLGEPBPATPAVGPPTMASGGGAPMADNNEGADGYNASGWNHCDSTLGRVIT 240
Db      181 SVDPQPLGEPBPATPAVGPPTMASGGGAPMADNNEGADGYNASGWNHCDSTLGRVIT 240
Qy      241 TTSTRTMALPTYNHLYKOISSASTGASNDNHYFGYSTPMWGFDPNRFCHFSRDMQRL 300
Db      241 TTSTRTMALPTYNHLYKOISSASTGASNDNHYFGYSTPMWGFDPNRFCHFSRDMQRL 299
Qy      301 INNNMGFRPKRLNFKLFNIOVKEVTTNDGVTTIANNLSTVQVFSDESQYLPYVLGSAHQ 360
Db      300 INNNMGFRPKRLNFKLFNIOVKEVTONGGTTIANNLSTVQVFTDSQYLPYVLGSAHQ 359
Qy      361 GCLPPPADVEMIPQYGYTLNNGSQAVGRSSFCLEYPFSPQMLRTGNNFTFSYTFEDVP 420
Db      360 GCLPPPADVEMIPQYGYTLNNGSQAVGRSSFCLEYPFSPQMLRTGNNFTFSYTFEDVP 419
Qy      421 FHSSTAHSGSLDRMLNPLIDQYLYLNTQONSGSAOKMDLFSRGPAGMSVOPKMWLP 480
Db      420 FHSSTAHSGSLDRMLNPLIDQYLYLNTQONSGSAOKMDLFSRGPAGMSVOPKMWLP 479
Qy      481 GPCYRQORVSKTKTNNNSNFTMGASKYNLNGRESIINPGTAMASHKDEDEKFPMSGV 540
Db      480 GPCYRQORVSKTSADNNNSSEYMTGATKYHLNGRDGLVNPBPAMASHKDEDEKFPMSGV 539
Qy      541 MIFGKSASASTALDNTVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAATGQVHAMG 600
Db      540 LIFGQKSEKIVNDIEKWMITDEBEIRITNPVATEOYGSVSTNLGRKQOATADAVNQG 599
Qy      601 ALPGWWMODRDVYLQGPIMAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTPVANPAP 660
Db      600 VLPGWWMODRDVYLQGPIMAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTPVANPAST 659
Qy      661 BBSATKFASTIOYSTGVSVIEWELQKNSKRNPEVQYTSNVAKSANVDFVNNGL 720
Db      660 TFSAKFASFTIOYSTGVSVIEWELQKNSKRNPEIQTYSNNKSNVNDFTVDTNGV 719
Qy      721 YTEPRPIGTRYLTRPL 736
Db      720 YSEPRPIGTRYLTRNL 735

RESULT 6
US-09-807-802A-15
; Sequence 15, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 599
; TYPE: PRT
; ORGANISM: AAV-1
```

```
US-09-807-802A-15
Query Match      81.5%; Score 3251; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 1,1e-276;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      138 TAPGKRPVBSPOEPDSSSIGIKTGOQPAKKRLNFGQTDSESVDPQPLGEPBPATPA 197
Db      1 TAPGKRPVBSPOEPDSSSIGIKTGOQPAKKRLNFGQTDSESVDPQPLGEPBPATPA 60
Qy      198 VGPPTMASGGGAPMADNNEGADGYNASGWNHCDSTLGRVITTTSTRTMALPTYNHLY 257
Db      61 VGPPTMASGGGAPMADNNEGADGYNASGWNHCDSTLGRVITTTSTRTMALPTYNHLY 120
Qy      258 KOISSASTGASNDNHYFGYSTPMWGFDPNRFCHFSRDMQRLINNNMGFRPKRLNFKLF 317
Db      121 KOISSASTGASNDNHYFGYSTPMWGFDPNRFCHFSRDMQRLINNNMGFRPKRLNFKLF 180
Qy      318 NIQVKEVTTNDGVTTIANNLSTVQVFSDESQYLPYVLGSAHQCLPPPADVEMIPQY 377
Db      181 NIQVKEVTTNDGVTTIANNLSTVQVFSDESQYLPYVLGSAHQCLPPPADVEMIPQY 240
Qy      378 YLTNNGSQAVGRSSFCLEYPFSPQMLRTGNNFTFSYTFEBVPHSSYAHSGSLDRMLNP 437
Db      241 YLTNNGSQAVGRSSFCLEYPFSPQMLRTGNNFTFSYTFEBVPHSSYAHSGSLDRMLNP 300
Qy      438 LIIDYLYLNTQONSGSAOKMDLFSRGPAGMSVOPKMWLPQPCYRQORVSKTKTDNN 497
Db      301 LIIDYLYLNTQONSGSAOKMDLFSRGPAGMSVOPKMWLPQPCYRQORVSKTKTDNN 360
Qy      498 NSNFTMGASKYNLNGRESIINPGTAMASHKDEDEKFPMSGVMI FKESASASTALDN 557
Db      361 NSNFTMGASKYNLNGRESIINPGTAMASHKDEDEKFPMSGVMI FKESASASTALDN 420
Qy      558 VMITDEBEIKATNPVATERFGTVAVNFQSSSTDPAATGQVHAMGLPGWWMODRDVYLQGP 617
Db      421 VMITDEBEIKATNPVATERFGTVAVNFQSSSTDPAATGQVHAMGLPGWWMODRDVYLQGP 480
Qy      618 IMAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTPVANPAPBBSATKFASTIOYSTG 677
Db      481 IMAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTPVANPAPBBSATKFASTIOYSTG 540
Qy      678 QVSVIEWELQKNSKRNPEVQYTSNVAKSANVDFVNNGLYTEPRPIGTRYLTRPL 736
Db      541 QVSVIEWELQKNSKRNPEVQYTSNVAKSANVDFVNNGLYTEPRPIGTRYLTRPL 599

RESULT 7
US-09-807-802A-17
; Sequence 17, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 534
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-17

Query Match      72.9%; Score 2906; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1,8e-246;
```

```
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 203 MASGGGAPMADNNEGADGVGNASGNMHCSTWLGDRVITTSRTTALPTYNHLYKQISS 262
DB 1 MASGGGAPMADNNEGADGVGNASGNMHCSTWLGDRVITTSRTTALPTYNHLYKQISS 60
QY 263 ASTGASNDNHRYFGYSTPMGWYFDPNRPFHCHFSPPDMQRLINNMGFPKRLNFKLNIQYK 322
DB 61 ASTGASNDNHRYFGYSTPMGWYFDPNRPFHCHFSPPDMQRLINNMGFPKRLNFKLNIQYK 120
QY 323 EYTTNDGVTITANNLITSTVOVPSDSRYQLPYVLGSHQGLPPFPADVMTIOYGLTLN 382
DB 121 EYTTNDGVTITANNLITSTVOVPSDSRYQLPYVLGSHQGLPPFPADVMTIOYGLTLN 180
QY 383 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPLIDQY 442
DB 181 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPLIDQY 240
QY 443 LYVLNRTQNGSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTKTDNNNSNFT 502
DB 241 LYVLNRTQNGSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTKTDNNNSNFT 300
QY 503 WTGASCTYNLNGRESITINPGTAMASHKDDKDFPMSGVMIFGKESAGSNTALDNMTD 562
DB 301 WTGASCTYNLNGRESITINPGTAMASHKDDKDFPMSGVMIFGKESAGSNTALDNMTD 360
QY 563 EEEIKATNPVATERFGTAVANFQSSSTDPATGVDHAMGALPGMWODRDVYLQGPIMAKI 622
DB 361 EEEIKATNPVATERFGTAVANFQSSSTDPATGVDHAMGALPGMWODRDVYLQGPIMAKI 420
QY 623 PHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPAPAEFSATKFASTTOYSTQVSYE 682
DB 421 PHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPAPAEFSATKFASTTOYSTQVSYE 480
QY 683 IEMELOKENSKRNPPEVOYTSTNYSKASANDFTVDNNGLYTERPPIGTRYLRPL 736
DB 481 IEMELOKENSKRNPPEVOYTSTNYSKASANDFTVDNNGLYTERPPIGTRYLRPL 534

RESULT 8
US-10-038-972A-14
; Sequence 14, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 2835/36960US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 598
; TYPE: PR
; ORGANISM: adeno-associated virus 2 VP2 capsid protein
US-10-038-972A-14

Query Match 69.2%; Score 2759.5; DB 2; Length 598;
Best Local Similarity 82.8%; Pred. No. 1,7e-233;
Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
```

```
DB 122 QISSGS -GASNDNHRYFGYSTPMGWYFDPNRPFHCHFSPPDMQRLINNMGFPKRLNFKLNI 180
QY 319 IQKEVTTNDGVTITANNLITSTVOVPSDSRYQLPYVLGSHQGLPPFPADVMTIOYGLTLN 378
DB 181 IQKEVTTNDGVTITANNLITSTVOVPSDSRYQLPYVLGSHQGLPPFPADVMTIOYGLTLN 240
QY 379 LTLNNGSGAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPL 438
DB 241 LTLNNGSGAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPL 300
QY 439 IDQVLYLNRTOQNGSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTKTDNNN 498
DB 301 IDQVLYLNRTOQNGSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTKTDNNN 360
QY 499 SNFTTASCTYNLNGRESITINPGTAMASHKDDKDFPMSGVMIFGKESAGSNTALDNV 558
DB 361 SEYSWTGATKYHLNGRDSTLVNPGPAMASHKDDKDFPMSGVMIFGKESAGSNTALDNV 420
QY 559 MTTDEEIKATNPVATERFGTAVANFQSSSTDPATGVDHAMGALPGMWODRDVYLQGP 618
DB 421 MTTDEEIKATNPVATERFGTAVANFQSSSTDPATGVDHAMGALPGMWODRDVYLQGP 480
QY 619 WAKIHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPAPAEFSATKFASTTOYSTQ 678
DB 481 WAKIHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPAPAEFSATKFASTTOYSTQ 540
QY 679 VSEIEMELOKENSKRNPPEVOYTSTNYSKASANDFTVDNNGLYTERPPIGTRYLRPL 736
DB 541 VSEIEMELOKENSKRNPPEVOYTSTNYSKASANDFTVDNNGLYTERPPIGTRYLRPL 598

RESULT 9
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chordini, John A.
; APPLICANT: Kolin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PR
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP1
US-09-532-594B-4

Query Match 62.3%; Score 2486.5; DB 2; Length 734;
Best Local Similarity 63.5%; Pred. No. 2.3e-209;
Matches 475; Conservative 81; Mismatches 161; Indels 31; Gaps 10;
```

Qy	183	FDDPPLSGPPRATPAVGVPTTMASSGGCAPMADNNBGADGVGNASGMMHCHSTYLGRVITTT	242
Db	183	PEGSTSG-----AMSDDSEMPAAAGAAVEGGQGDGVGNASGDWCHDSTYSBGRVITTT	236
Qy	243	STRTALPTVNNHLLKQJISASSTAGSANDNHVGYSTPMGYPDENPFHCFSPRDMQRLIN	302
Db	237	STRITVLPPTVNNHLLKRLGE-----SLQSTVTVGFSTPMGYPFENPFHCFSPRDMQRLIN	292
Qy	303	NNMGFRPRRLNFKLFNIOYKEVTTNDGVTTLANLITSTVOVPSDSEYQLPYVLGSAHQCC	362
Db	293	NNMGFRPRRLNFKLFNIOYKEVTTNGETTVANNLITSTVOIPADSSYELPYVMDAQEGSS	352
Qy	363	LPRPADVPMIPOYCY---LTLNNSQAVGRSSFCLEBFPBOMLRTGNPFSTFEEVY	419
Db	353	LPRPNDVPMVPOYCGVCGLVGTNTSQOQTRNAFCLEBFPBOMLRTGNPFITTSFEVY	412
Qy	420	PFHSSVYASQSILDRMLNPLIDOLYLYLNTTON---QSGSAONKDLLFSGSPAGMSVOP	475
Db	413	PFHSHVYASQSILDRMLNPLIDOLYLMGLQSTTGTTLNAQTATNT---PFLKLNPTNSNFK	469
Qy	476	KNMLFGPCYRQGRVSKTKTDNNNSNFTWTGAS---KY-----NLNGRESLIINPTAMASHK	528
Db	470	KNMLFGPSPKIQGQFSPKXTA--NONYKIPATGSSDLIKYETHSTLDGWSALTEGPPMATAG	527
Qy	529	DDEDFFPMSGMIMGKESAGASNTALDNVMTIDEEIKATPVALTERGTVAVUNQSSS	588
Db	528	PADSK--FNSNSQLFPGPPKONGNTATVPGTLITSEBELAATNATDTDMGNLPGCDQSSS	586
Qy	589	TDPATGDVHAMGALPGMWWQDRDVYLOGPIMAKIPHTDGHFSPSLMGSGFGLNPPPIIL	648
Db	587	NLPVTVDRLLTALGAVPGMWWQNDIYVGGPIMAKIPHTDGHFSPSLIGGFGGLNPPPIIF	646
Qy	649	IKNTVVPANPRPEFSAITKASTITTOYSTQGVSEIEMELQKENSXKMNBEVOTTSYAAKS	708
Db	647	IKNTVVPANPATTFSTPNSFITOYSTQGVSOVDMELOKERSKXMNBEVOTTSYVGOQ	706
Qy	709	ANVDFTVDNNGLYTEPRPIGRYLRPL	736
Db	707	NSLLMAPDAAGKYTEPRALGRYLRTHL	734
RESULT 10			
US-10-038-972A-15			
Sequence 15, Application US/100389972A			
Patent No. 6962815			
GENERAL INFORMATION:			
APPLICANT: J. Bartlett			
TITLE OF INVENTION: AAV VECTORS AND METHODS			
FILE REFERENCE: 28335/3696US			
CURRENT APPLICATION NUMBER: US/10/038,972A			
CURRENT FILING DATE: 2002-01-04			
PRIOR APPLICATION NUMBER: US 60/260,124			
PRIOR FILING DATE: 2001-01-05			
NUMBER OF SEQ ID NOS: 18			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 15			
LENGTH: 533			
TYPE: PRT			
ORGANISM: adeno-associated virus 2 VP3 capsid protein			
US-10-038-972A-15			

Query Match	62.2%	Score 2481.5	DB 2	Length 533
Best Local Similarity	83.3%	Pred. No. 3.7e-209		
Matches 445	Conservative 37	Mismatches 51	Indels 1	Gaps 1
QY	203	MASGGCAPMADNNNEGADGVGNASGNWCHSTVLGDRVITTSRTMALPTYYNNHLKQISS	262	
Db	1	MATSGCAPMADNNEGADGVGNSSGNWCHSTVMGDRVITTSRTMALPTYYNNHLKQISS	60	
QY	263	ASTGASNDNHRYGYSTPWPGYPDFNRRHCHFSRPDMQRLINNWWGFPKRLNRLFLPIQYK	322	
Db	61	GS- GASNDNHRYGYSTPWPGYPDFNRRHCHFSRPDMQRLINNWWGFPKRLNRLFLPIQYK	119	

QY	323	IVTTNNDDVTTTANNLTSTVQVSSDBEYOLPYVLGSHQOCLPPFPADVMIMEPYGVLTLN	382
Db	120	EYTONDGTITTTIANNLTSTVQVFTDBEYOLPYVLGSHQOCLPPFPADVMIMEPYGVLTLN	179
QY	383	NGSQAVGRSSFYCLEYFPSPQMLRTNNFTFSYTFEEVPFHSASHAQSLDRLLNPLIDY	442
Db	180	NGSQAVGRSSFYCLEYFPSPQMLRTNNFTFSYTFEEVPFHSASHAQSLDRLLNPLIDY	239
QY	443	LYYLNRTONQSGSAONKDLFBSRGSYPAGMSVOPKXWLPBPCTRQORVSKTKTDNNNSFT	502
Db	240	LYYLSTRTPSGTITTTQSRLOFSGAASDIRDQSRMWLPBPCROQGVSKTSADNNNSYS	299
QY	503	WTGASKNUNLGESEIINPGTAMASHKHDBDKFPSPGVMI FGSASGANTLADNMITD	562
Db	300	WTGATKYHLNGDSTLVNPGPAMASHKHDBDKFPSPGVMI FGSQSEKTNVDEKXMITD	359
QY	563	EEBITATPVATERFGTVAVNFQSSSTDATGDVHAMGALPGVMQODRDVYLOGPIWAKI	622
Db	360	EEBITTPVPALREOYGVSVSTNLQDRGNROAAITADVNTQVLPQVMQODRDVYLOGPIWAKI	419
QY	623	PHTHGHFHPSPIMGFGFGLKNPPOPILLIKNTPVANPAPAEFSATKEAFTTQYSTGVSVF	682
Db	420	PHTHGHFHPSPIMGFGFGLKHPPOLLIKNTPVANPASTFESAKFASFTQYSTGVSVF	479
QY	683	IMELJOKENSKRNBEVOYTSNYAKSANDFTVDNNGUYTERPPIGSTRVLTREL	736
Db	480	IMELJOKENSKRNBEIOYTSNYKNSVANDFTVDNNGVSEBPPIGSTRVLTREL	533

```

RESULT 11
US-09-533-427-4
  / Sequence 4, Application US/09533427
  / Patent No. 6855314
  / GENERAL INFORMATION:
  / APPLICANT: Chiorini, John
  / APPLICANT: Kotlin, Robert M.
  / APPLICANT: Safer, Brian
  / APPLICANT: Davidson, Elizabeth
  / APPLICANT: Zahner, Joseph
  / TITLE OF INVENTION: AAV5 VECTOR FOR TRANSDUCING BRAIN CELLS AND LONG CELLS
  / FILE REFERENCE: 14014.03302
  / CURRENT APPLICATION NUMBER: US/09/533,427
  / CURRENT FILING DATE: 2000-03-22
  / NUMBER OF SEQ ID NOS: 23
  / SOFTWARE: FastSeq for Windows Version 3.0
  / SEQ ID NO 4
  / LENGTH: 724
  / TYPE: PRT
  / ORGANISM: Artificial Sequence
  / FEATURE:
  / OTHER INFORMATION: Description of Artificial Sequence: /No. 6855314e =
  / OTHER INFORMATION: synthetic construct
US-09-533-427-4

```

[illegible]

Db	9	AAAGGAAGV-BGGGAGDGVGNASGDHMDCTWMSGHVTTTSRTTMYLPTYNHLYKRLGE-	66
Qy	264	STGASNDNHFGYSTWGWYDFENRFCHFSPRDMORLLNNNGRRPRKLFKLPNIQVE	323
Db	67	---LSQSTNTYNGFSTPGYDFNRFCHFSPRDMORLLNNNGRRPRKLFKLPNIQVE	123
Qy	324	VTNNQDVTTLANNLTSTVGVFSDSEYOLPVLDSAHQGCCLPPRAVDFMIPOYGV---LT	380
Db	124	VTTNSNGTITVANNLTSTVQIFADSYELPYVMDAGOGSLPPRPNDVFMVPOYGCGLVT	183
Qy	381	LNNGSQAVGRSSFCLEYPFSQMLRTGNNTFFSYEEVEFHSYAHOSLDRMLNPLID	440
Db	184	GNISQOOTDNNATYCLEYFSPQMLRTGNNTFFSYEEVEFHSYAHOSLDRMLNPLID	243
Qy	441	QYLYVYNNRTON---QSGSAQNKDILFSRGSPPAGMSVOPQWMLPGCYTCROORVSKTIDN	496
Db	244	QYLMGLOSTTGTTLNAGTATTN---FTKLRPTNFSAFKIXMWLGPSPILKQOGFEKTA--N	298
Qy	497	NNSFTWTTGAS---KY---NLNGRBSIINPGTAMASHKODEDFPMGSMIFGKXSAG	549
Db	299	QNKYIPATYGDLSLIKETHSTLDGRMALTPGPMAIAGPADSK-FSNSQLIFAGPHONG	357
Qy	550	ASNTALDNWITDEEBEIKATNPVATERFGTVAVFQSSSTDPAIGDVAHAGALPGWYOD	609
Db	358	NTATVPETILLFTSBEELATINADITDWMQWLPQGDQSNLSLPTYDRLTALGAVGWYQWON	417
Qy	610	RDVYLQCPIMAKIPHTDGHFHPSPDLMGFGKLPNPOLIKNTPVANPPAEPGATFPAS	669
Db	418	RDIYYQGPIMAKIPHTDGHFHPSPDLMGFGKLPNPOLIKNTPVANPPAEPGATFPAS	477
Qy	670	PIQVSTNGOVSVELEWELQKENSRRNPDEVQYSNVKASANVDTVTNNGLYTPRPDGT	729
Db	478	FIQVSTNGOVSVOYDWEIQKERSRRNPDEVQYSNVQOQNSLWAPDAAGKTEPPRAIGT	537
Qy	730	RYLTRPL 736	
Db	538	RYLTRHL 544	
RESULT 15			
US-09-533-427-6			
Sequence 6, Application US/09533427			
Patent No. 685314			
GENERAL INFORMATION:			
APPLICANT: Chiorini, John			
APPLICANT: Kotlin, Robert M.			
APPLICANT: Saefer, Brian			
APPLICANT: Davidson, Elizabeth			
TITLE OF INVENTION: AAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS			
FILE REFERENCE: 14014_032302			
CURRENT APPLICATION NUMBER: US/09/533,427			
CURRENT FILING DATE: 2000-03-22			
NUMBER OF SEQ ID NOS: 23			
SOFTWARE: FastSeq for Windows Version 3.0			
SEQ ID NO 6			
LENGTH: 532			
TYPE: PRT			
ORGANISM: Artificial Sequence			
FEATURE:			
OTHER INFORMATION: Description of Artificial Sequence:/No. 685314e =			
US-09-533-427-6			
Query Match			
Best Local Similarity 58.9%; Pred. No. 2,1e-137; Indels 14; Gaps 5;			
Matches 318; Conservative 54; Mismatches 154;			
Qy	203	MASGGGAPMAIDNNGADGVGNASGNWHCDSTMLCDRYLTSTRTMALPTYNHLYKQIIS	262
Db	1	MSAGGGPFLGDNNNGADGVGNASGDHMDCTWMDRVTYVSTRTMYLPSYNNHQYREIKS	60
Qy	263	ASTGASDNHNYFGYSTWGWYDFENRFCHFSPRDMORLLNNNGRRPRKLFKLPNIQVK	322

Db	61	GSVDSGNANAYFGYSTPMCGYFDPNRPHSHWSRDMORLNNYWGFRPRSLRYKIFNIQVK	120
Qy	323	EYTTNDGYTTIANNTSTVOVPSDSEYQLPYVLGSANHOCLEPPADVEMIDPOXYLTIN	382
Db	121	EYTVQDSTTTIANNLSTVOVFPTDDYQUPYVNGNTEGCLPAFPQVFTLPLQYGVATLN	180
Qy	383	--NGSQAVGRSGSFCLFYFSPQMLRTGNNFTSYTFEYVPHSHSYAHQSOLRLMPLID	440
Db	181	RDNTENPERSSFPCLFYEPSPKMLRTGNNFEFTYNEEVPFHSSFPASQMLPLDANPLVD	240
Qy	441	OYLXYLNTQONGSGANOKDLEFSRGSFAGMSQOP-KQMLPGCCVYQQRSVSKTKTDNNNS	499
Db	241	OYLXYFVSNTNNGVOFNKLN-----AGRANYTKMFPGBMGHTQCMNLGCVNRAS	293
Qy	500	NFTWAGASKYNTNGRESIINPCTAMASHDODEKFPFPGMGVMI FGSBSAGASNTAL---	556
Db	294	VSAPFTTRMELFGASYOUPPOPGNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLBG	353
Qy	557	NVMTIDDEEIKATNPVATERFCTVAVNFQSSSTDPATGVDHANGALPGMYQODRVYLOQ	616
Db	354	NMLITSESETPGVNRVAYVVGQMAFNNOGSTPAATGYTNLOEIPGVSWMERDVLQO	413
Qy	617	PIWAKIPIHTDGFHPHSPPLMGFGRLKNRPPQILIKTVPVANPAAESATKFAFIIOYST	676
Db	414	PIWAKIPELGAHFHPHSPAMGGFOLKTRPPMLIKTNPVPRN-ITSPSDVPVSSFTIOYST	472
Qy	677	GOVSVEIEMELKENSKRNNPEYOQYSNYSKASANDVFTVDNNLCLYNEPRPIGRYTYRPL	736
Db	473	GOVYTEMEMELKENSKRNNPELOYNNTNNNDPOVDPAEDSTSEYITTRPIGRYTYRPL	532

Search completed: November 23, 2005, 17:42:46
Job time : 36.4724 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:25:34 / Search time 141.766 Seconds
(without alignment)
3662.868 Million cell updates/sec

Title: US-10-696-282-13
Perfect score: 3989
Sequence: 1 MAA0GYPWLEMDNLSEGR.....NNGLYTEPRPIGTRYLRPL 736

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_05.80:*
2: uniprot_eprot:*
3: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Rank	Score	Query	Match	Length	DB	ID	Description
1	3989	100.0	736	2	Q9WB88	SVIRU	Q9WB88 adeno-ssoc
2	3968	99.5	736	2	Q6JC08	SVIRU	Q6JC08 adeno-ssoc
3	3963	99.3	736	2	Q56137	SVIRU	Q56137 adeno-ssoc
4	3883.5	97.4	737	2	Q6JC13	SVIRU	Q6JC13 adeno-ssoc
5	3873	97.1	736	2	Q6JC12	SVIRU	Q6JC12 adeno-ssoc
6	3853	96.6	736	2	Q6JC10	SVIRU	Q6JC10 adeno-ssoc
7	3511	88.0	736	2	Q56139	SVIRU	Q56139 adeno-ssoc
8	3494	87.6	736	2	Q65311	SVIRU	Q65311 adeno-ssoc
9	3473	87.1	736	2	Q67008	SVIRU	Q67008 adeno-ssoc
10	3469.5	87.0	735	2	Q6JB27	SVIRU	Q6JB27 adeno-ssoc
11	3456.5	86.7	737	2	Q6JC52	SVIRU	Q6JC52 adeno-ssoc
12	3456.5	86.6	737	2	Q6JC58	SVIRU	Q6JC58 adeno-ssoc
13	3455.5	86.6	737	2	Q6JC28	SVIRU	Q6JC28 adeno-ssoc
14	3453.5	86.6	735	2	Q6JC19	SVIRU	Q6JC19 adeno-ssoc
15	3453	86.6	738	2	Q6JC17	SVIRU	Q6JC17 adeno-ssoc
16	3451.5	86.5	735	2	Q6JC34	SVIRU	Q6JC34 adeno-ssoc
17	3450.5	86.5	737	2	Q6JC51	SVIRU	Q6JC51 adeno-ssoc
18	3449.5	86.5	735	2	Q6JB21	SVIRU	Q6JB21 adeno-ssoc
19	3449.5	86.5	735	2	Q6JC42	SVIRU	Q6JC42 adeno-ssoc
20	3449	86.5	738	2	Q5Y9B4	SVIRU	Q5Y9B4 adeno-ssoc
21	3449	86.5	738	2	Q6JC37	SVIRU	Q6JC37 adeno-ssoc
22	3448	86.4	738	2	Q808W5	SVIRU	Q808W5 non-human p
23	3446.5	86.4	735	2	Q6JC17	SVIRU	Q6JC17 adeno-ssoc
24	3445	86.4	738	2	Q6JC60	SVIRU	Q6JC60 adeno-ssoc
25	3443.5	86.3	737	2	Q6JC46	SVIRU	Q6JC46 adeno-ssoc
26	3442.5	86.3	735	2	Q6JC36	SVIRU	Q6JC36 adeno-ssoc
27	3442.5	86.3	737	2	Q8JCG0	SVIRU	Q8JCG0 adeno-ssoc
28	3442	86.3	738	2	Q6JC44	SVIRU	Q6JC44 adeno-ssoc
29	3442	86.3	738	2	Q6JC46	SVIRU	Q6JC46 adeno-ssoc
30	3441.5	86.3	735	2	Q6JC44	SVIRU	Q6JC44 adeno-ssoc
31	3440	86.2	738	2	Q6JC15	SVIRU	Q6JC15 adeno-ssoc

32	3440	86.2	738	2	Q6JC16	SVIRU	Q6JC16 adeno-ssoc
33	3440	86.2	738	2	Q6JC62	SVIRU	Q6JC62 adeno-ssoc
34	3439	86.2	738	2	Q6JB22	SVIRU	Q6JB22 adeno-ssoc
35	3438	86.2	734	2	Q6JC02	SVIRU	Q6JC02 adeno-ssoc
36	3438	86.2	738	2	Q808X3	SVIRU	Q808X3 non-human p
37	3437	86.2	738	2	Q6JC47	SVIRU	Q6JC47 adeno-ssoc
38	3437	86.2	738	2	Q6JC61	SVIRU	Q6JC61 adeno-ssoc
39	3436	86.1	734	2	Q6JC04	SVIRU	Q6JC04 adeno-ssoc
40	3434.5	86.1	735	2	Q6JB26	SVIRU	Q6JB26 adeno-ssoc
41	3434	86.1	728	2	Q808X4	SVIRU	Q808X4 non-human p
42	3434	86.1	738	2	Q6JB23	SVIRU	Q6JB23 adeno-ssoc
43	3433	86.1	738	2	Q6JC49	SVIRU	Q6JC49 adeno-ssoc
44	3433	86.1	738	2	Q6JC53	SVIRU	Q6JC53 adeno-ssoc
45	3432	86.0	738	2	Q6JC54	SVIRU	Q6JC54 adeno-ssoc

ALIGNMENTS

RESULT 1
Q9WB88 SVIRU PRELIMINARY; PRT; 736 AA.
AC Q9WB88;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=85106;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9214338; PubMed=10196295;
RA Xiao W., Chirmule N., Barta S.C., McCullough B., Gao G., Wilson J.M.;
RT "Gene therapy vectors based on adeno-associated virus type 1."
RL J. Virol. 73:3994-4003(1999).
DR EMBL; AF063497; A027757.1; -; Genomic_DNA.
DR SMR; Q9WB88; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; P:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo coat; 1.
SQ SEQUENCE 736 AA; 81375 MW; CPABFB9BDSOD0595 CRC64;

Query Match 100.0%; Score 3989; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.1e-249;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAA0GYPWLEMDNLSEGRWMDLKPAPKPAKNOQKDDGGLVLPQYKYLGPNGLD 60
1 KCEPVAAALAEHDKAVDOOLKAGDNPYLKRNHDAFEORLEQDTSFGNLAGAVRQ 120
61 KCEPVAAALAEHDKAVDOOLKAGDNPYLKRNHDAFEORLEQDTSFGNLAGAVRQ 120
1 MAA0GYPWLEMDNLSEGRWMDLKPAPKPAKNOQKDDGGLVLPQYKYLGPNGLD 60
61 KCEPVAAALAEHDKAVDOOLKAGDNPYLKRNHDAFEORLEQDTSFGNLAGAVRQ 120
121 AKKRVLEPLGLVEGAKTAPGKKRPVQSGPOEBSGGIGKTGQOPAKKRLNFGQTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKKRPVQSGPOEBSGGIGKTGQOPAKKRLNFGQTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKKRPVQSGPOEBSGGIGKTGQOPAKKRLNFGQTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKKRPVQSGPOEBSGGIGKTGQOPAKKRLNFGQTGDS 180
181 SVDPDPPLEPAPTPAAGPTTMASSGGAAPMADNNGACGCGNAGSNMHCDSITWLDLDRVI 240
181 SVDPDPPLEPAPTPAAGPTTMASSGGAAPMADNNGACGCGNAGSNMHCDSITWLDLDRVI 240
181 SVDPDPPLEPAPTPAAGPTTMASSGGAAPMADNNGACGCGNAGSNMHCDSITWLDLDRVI 240
241 TTSRTFWALPTYNNHLVKOISSASTGASNDNHYFGYSTPMGYDFNRFHCFSPRDMQRL 300
241 TTSRTFWALPTYNNHLVKOISSASTGASNDNHYFGYSTPMGYDFNRFHCFSPRDMQRL 300
241 TTSRTFWALPTYNNHLVKOISSASTGASNDNHYFGYSTPMGYDFNRFHCFSPRDMQRL 300
301 INNNWGFRRRLNFKLNFIVKEVTTNDGVTIANNLSTVOYFSDSEYQLPVYLSAQH 360
301 INNNWGFRRRLNFKLNFIVKEVTTNDGVTIANNLSTVOYFSDSEYQLPVYLSAQH 360
301 INNNWGFRRRLNFKLNFIVKEVTTNDGVTIANNLSTVOYFSDSEYQLPVYLSAQH 360
361 GCLPPRPADVFMIPQYGYLTINNGSQAVGRSSSYCLEFPFSQWLRTGNNTFTFYTEEVP 420

```
Db 361 GCLPPFADVFMIPQYGLTLNNGSQA VGRSSFYCLEYFP SQMLRTGNNFTFSYTFEEVP 420
Qy 421 FHSYASHQSIDLRLMNPILIDQYLYLNRTONQSSAQNKDLLFSGSPAGMSVQPKWLP 480
Db 421 FHSYASHQSIDLRLMNPILIDQYLYLNRTONQSSAQNKDLLFSGSPAGMSVQPKWLP 480
Qy 481 GPCYRQOVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDKFFPMSCV 540
Db 481 GPCYRQOVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDKFFPMSCV 540
Qy 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMG 600
Db 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMG 600
Qy 601 ALPGWVQDRDVLVYLGPIWAKIPHTDGHFHSPLMGSGFGLKNPPQILLIKTTPVAPNPA 660
Db 601 ALPGWVQDRDVLVYLGPIWAKIPHTDGHFHSPLMGSGFGLKNPPQILLIKTTPVAPNPA 660
Qy 661 EFSATKFPASFTIOYSTGVSVIEIMELQENSKRMNPEVQYTSNYAKSANDFTVDNGL 720
Db 661 EFSATKFPASFTIOYSTGVSVIEIMELQENSKRMNPEVQYTSNYAKSANDFTVDNGL 720
Qy 721 YTEBRPIGTRYLTRPL 736
Db 721 YTEBRPIGTRYLTRPL 736

RESULT 2
06JUC08 9VIRU PRELIMINARY; PRT; 736 AA.
ID 06JUC08
AC 06JUC08
DT 05-JUN-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Dependovirus.
OC NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL: AY530611; AAS99296.1; -; Genomic_DNA.
DR SMR: 06JUC08: 217-736.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81317 MW; 373C7C40975CD9BD CRC64;

Query Match 99.5%; Score 3968; DB 2; Length 736;
Best Local Similarity 99.3%; Pred. No. 2,6e-248;
Matches 731; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
Db 181 SVDPDQPLGEPPATPAVGFPTMASGGA PMADNNEGADGVNAGSNHCHSTWLGDRVI 240
Qy 241 TTSRTALPTYNHLYKQISSASTGASNDNHRYGYSYTPMGYDFPNFCHFSRDMQRL 300
Db 241 TTSRTALPTYNHLYKQISSASTGASNDNHRYGYSYTPMGYDFPNFCHFSRDMQRL 300
Qy 301 INNMGRPKRLNKLNIQVEYTTNDGYTTIANNTSTVQVSDSEYQLPYLGSAAHQ 360
Db 301 INNMGRPKRLNKLNIQVEYTTNDGYTTIANNTSTVQVSDSEYQLPYLGSAAHQ 360
Qy 361 GCLPPFADVFMIPQYGLTLNNGSQA VGRSSFYCLEYFP SQMLRTGNNFTFSYTFEEVP 420
Db 361 GCLPPFADVFMIPQYGLTLNNGSQA VGRSSFYCLEYFP SQMLRTGNNFTFSYTFEEVP 420
Qy 421 FHSYASHQSIDLRLMNPILIDQYLYLNRTONQSSAQNKDLLFSGSPAGMSVQPKWLP 480
Db 421 FHSYASHQSIDLRLMNPILIDQYLYLNRTONQSSAQNKDLLFSGSPAGMSVQPKWLP 480
Qy 481 GPCYRQOVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDKFFPMSCV 540
Db 481 GPCYRQOVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDKFFPMSCV 540
Qy 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMG 600
Db 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMG 600
Qy 601 ALPGWVQDRDVLVYLGPIWAKIPHTDGHFHSPLMGSGFGLKNPPQILLIKTTPVAPNPA 660
Db 601 ALPGWVQDRDVLVYLGPIWAKIPHTDGHFHSPLMGSGFGLKNPPQILLIKTTPVAPNPA 660
Qy 661 EFSATKFPASFTIOYSTGVSVIEIMELQENSKRMNPEVQYTSNYAKSANDFTVDNGL 720
Db 661 EFSATKFPASFTIOYSTGVSVIEIMELQENSKRMNPEVQYTSNYAKSANDFTVDNGL 720
Qy 721 YTEBRPIGTRYLTRPL 736
Db 721 YTEBRPIGTRYLTRPL 736

RESULT 3
056137 9VIRU PRELIMINARY; PRT; 736 AA.
ID 056137
AC 056137
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus 6.
OC Viruses; ssDNA viruses; Parvoviridae; Dependovirus.
OC NCBI_TaxID=68558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE:98080418; PubMed:9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
RT (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
DR EMBL: AF028704; AAB95450.1; -; Genomic_DNA.
DR SMR: 056137: 217-736.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;

Query Match 99.3%; Score 3963; DB 2; Length 736;
Best Local Similarity 99.2%; Pred. No. 5,4e-248;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

QY 61 KEEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADADEFORLEQDTSFGGNLGRAVQ 120
DB 61 KEEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADADEFORLEQDTSFGGNLGRAVQ 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQDGS 180
DB 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQDGS 180
QY 181 SVDPDQPLGEPAPATPAVPTTMASSGGAPMADNNEGADGVGNASGNWCHDSTWLGDYV 240
DB 181 SVDPDQPLGEPAPATPAVPTTMASSGGAPMADNNEGADGVGNASGNWCHDSTWLGDYV 240
QY 241 TTSTRTALPTYNNHLVKQISSASTGASNDNHFGYSTPWGYPFENRPHCHSPRDMQR 300
DB 241 TTSTRTALPTYNNHLVKQISSASTGASNDNHFGYSTPWGYPFENRPHCHSPRDMQR 300
QY 301 INNMGFRPKRLNFKLFINIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
DB 301 INNMGFRPKRLNFKLFINIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
QY 361 GCLPPPADVFMIPQGYLTLLNGSOAVGRSSFYCLEYFPSSOMLRTGNNFTFSYTFEEV 420
DB 361 GCLPPPADVFMIPQGYLTLLNGSOAVGRSSFYCLEYFPSSOMLRTGNNFTFSYTFEEV 420
QY 421 FHSSYAHOSQDLRLNPLIDQYLVLNRTQNGSAGQNDLFSRGSFAGMSVQPKWLP 480
DB 421 FHSSYAHOSQDLRLNPLIDQYLVLNRTQNGSAGQNDLFSRGSFAGMSVQPKWLP 480
QY 481 GPCYRQORVSKTIDNNNSFTWTGASKYLNLRRESITNGTMAHSHKDEDEKFFPMG 540
DB 481 GPCYRQORVSKTIDNNNSFTWTGASKYLNLRRESITNGTMAHSHKDEDEKFFPMG 540
QY 541 MIFGKESAGASNTALDNMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGVDHAM 600
DB 541 MIFGKESAGASNTALDNMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGVDHAM 600
QY 601 ALPGMWODRDVYLOGPIWAKIPHTDGHFHPSPLMGFGKLPPOILLKNTVPANPP 660
DB 601 ALPGMWODRDVYLOGPIWAKIPHTDGHFHPSPLMGFGKLPPOILLKNTVPANPP 660
QY 661 EFSATKPAFTTQYSTGVQSVIEIEMELQKNSKRNPEVQYTSNYAKSANVDFTVNNG 720
DB 661 EFSATKPAFTTQYSTGVQSVIEIEMELQKNSKRNPEVQYTSNYAKSANVDFTVNNG 720
QY 721 YTEPRPIGRYLRPL 736
DB 721 YTEPRPIGRYLRPL 736

DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
DR SEQUENCE 737 AA; 81332 MW; 94CED0F3AFD6E492 CRC64;
Query Match 97.4%; Score 3883.5; DB 2; Length 737;
Best Local Similarity 97.4%; Pred. No. 7.7e-243;
Matches 718; Conservative 6; Mismatches 12; Indels 1; Gaps 1;
QY 1 MAADGYLPDWEEDNLSEGIREWMDLKPQAPKPKANQOKODDGRGLVPEYKYLGPENGLD 60
DB 1 MAADGYLPDWEEDNLSEGIREWMDLKPQAPKPKANQOKODDGRGLVPEYKYLGPENGLD 60
QY 61 KEEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADADEFORLEQDTSFGGNLGRAVQ 120
DB 61 KEEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADADEFORLEQDTSFGGNLGRAVQ 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQDGS 179
DB 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQDGS 180
QY 180 ESVDPDQPLGEPAPATPAVPTTMASSGGAPMADNNEGADGVGNASGNWCHDSTWLGDYV 239
DB 181 ESVDPDQPLGEPAPATPAVPTTMASSGGAPMADNNEGADGVGNASGNWCHDSTWLGDYV 240
QY 240 TTSTRTALPTYNNHLVKQISSASTGASNDNHFGYSTPWGYPFENRPHCHSPRDMQR 239
DB 241 TTSTRTALPTYNNHLVKQISSASTGASNDNHFGYSTPWGYPFENRPHCHSPRDMQR 300
QY 300 LNNMGFRPKRLNFKLFINIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 359
DB 301 LNNMGFRPKRLNFKLFINIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
QY 360 GCLPPPADVFMIPQGYLTLLNGSOAVGRSSFYCLEYFPSSOMLRTGNNFTFSYTFEEV 419
DB 361 GCLPPPADVFMIPQGYLTLLNGSOAVGRSSFYCLEYFPSSOMLRTGNNFTFSYTFEEV 420
QY 420 PHSSYAHOSQDLRLNPLIDQYLVLNRTQNGSAGQNDLFSRGSFAGMSVQPKWLP 479
DB 421 PHSSYAHOSQDLRLNPLIDQYLVLNRTQNGSAGQNDLFSRGSFAGMSVQPKWLP 480
QY 480 GPCYRQORVSKTIDNNNSFTWTGASKYLNLRRESITNGTMAHSHKDEDEKFFPMG 539
DB 481 GPCYRQORVSKTIDNNNSFTWTGASKYLNLRRESITNGTMAHSHKDEDEKFFPMG 540
QY 540 VMIFGKESAGASNTALDNMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGVDHAM 599
DB 541 VMIFGKESAGASNTALDNMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGVDHAM 600
QY 600 GALPGMWODRDVYLOGPIWAKIPHTDGHFHPSPLMGFGKLPPOILLKNTVPANPP 659
DB 601 GALPGMWODRDVYLOGPIWAKIPHTDGHFHPSPLMGFGKLPPOILLKNTVPANPP 660
QY 660 AFSATKPAFTTQYSTGVQSVIEIEMELQKNSKRNPEVQYTSNYAKSANVDFTVNNG 719
DB 661 AFSATKPAFTTQYSTGVQSVIEIEMELQKNSKRNPEVQYTSNYAKSANVDFTVNNG 720
QY 720 LYTEPRPIGRYLRPL 736
DB 721 LYTEPRPIGRYLRPL 737

RESULT 4
ID 06JC13_9VIRU PRELIMINARY; PRT; 737 AA.
AC 06JC13;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues".
RL J. Virol. 78:6381-6388(2004).
DR EMBL, AY530606; M599291.1; -, Genomic_DNA.
DR SMR; 06JC13; 218-737.
DR GO; GO:0019028; C:viral capsid; IEA.

RESULT 5
ID 06JC12_9VIRU PRELIMINARY; PRT; 736 AA.
AC 06JC12;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

OX NCBI_TaxID=272636;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
 RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
 RT "Clades of Adeno-associated viruses are widely disseminated in human
 RT tissues.";
 RL J. Virol. 78:6381-6388(2004).
 DR EMBL: AY530607; AAS9292.1; -, Genomic_DNA.
 DR SMR: Q6JC12; 217-736.
 DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 736 AA; 81364 MW; F2D75A06662E6376 CRC64;

Query Match 97.1%; Score 3873; DB 2; Length 736;
 Best Local Similarity 96.9%; Pred. No. 3.7e-242;
 Matches 713; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

OY	1	MAADGVLDMLENDLSGIRREWDLKPGAPKPKANQKODDORGLVLPGYKYLGPFGND	60
DB	1	MAADGVLDMLENDLSGIRREWDLKPGAPKPKANQKODDORGLVLPGYKYLGPFGND	60
OY	61	KGEFVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGGNGRAVFO	120
DB	61	KGEFVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGGNGRAVFO	120
OY	121	AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSGIGTKGQOPAKKRLNFGQTGDS	180
DB	121	AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSGIGTKGQOPAKKRLNFGQTGDS	180
OY	181	SVPDPQPLGEPAPATPAAGPTTMAAGGAPMADNNEGADGVNAGNHCDSITWLGDRVI	240
DB	181	SVPDPQPLGEPAPATPAAGPTTMAAGGAPMADNNEGADGVNAGNHCDSITWLGDRVI	240
OY	241	TTSTRTMAALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDNRRHCHFSPPDMQRL	300
DB	241	TTSTRTMAALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDNRRHCHFSPPDMQRL	300
OY	301	INNMGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQ	360
DB	301	INNMGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQ	360
OY	361	GCLPPFPADVPMIPOYGITLTNNGSOAVGRSSFCLEFPQMLRTGNFTFSYFEEVP	420
DB	361	GCLPPFPADVPMIPOYGITLTNNGSOAVGRSSFCLEFPQMLRTGNFTFSYFEEVP	420
OY	421	FHSSVAHSOSLDRLNMPILIDQYLYLTNRTONQSGAONKDLFSRGSPPAGMSVOPKMLP	480
DB	421	FHSSVAHSOSLDRLNMPILIDQYLYLTNRTONQSGAONKDLFSRGSPPAGMSVOPKMLP	480
OY	481	GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMGCV	540
DB	481	GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMGCV	540
OY	541	MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPATGTVHAMG	600
DB	541	MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPATGTVHAMG	600
OY	601	ALPGWVQORVDYVLOQPIWAKIPIHTDGHFHSPLMGFGELKNPPOILIKNTVPANPA	660
DB	601	ALPGWVQORVDYVLOQPIWAKIPIHTDGHFHSPLMGFGELKNPPOILIKNTVPANPA	660
OY	661	EFSATKFAAFITQYSTGVSVLEIWELOKENSKNMPEVOYTSNKAANVDFTVDNGL	720
DB	661	EFSATKFAAFITQYSTGVSVLEIWELOKENSKNMPEVOYTSNKAANVDFTVDNGL	720
OY	721	YTEPRPIGTRYLTRPL 736	
DB	721	YTEPRPIGTRYLTRPL 736	

RESULT 6
 O6JC10 VIRUS PRELIMINARY; PRT; 736 AA.
 AC O6JC10;
 DT 05-JUL-2004 (Tremblere, 27, Created)
 DT 05-JUL-2004 (Tremblere, 27, Last sequence update)
 DT 05-JUL-2004 (Tremblere, 27, Last annotation update)
 DE Capsid protein VP1.
 GN Name=cap;
 OS Adeno-associated virus.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
 OX NCBI_TaxID=272636;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
 RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
 RT "Clades of Adeno-associated viruses are widely disseminated in human
 RT tissues.";
 RL J. Virol. 78:6381-6388(2004).
 DR EMBL: AY530609; AAS9294.1; -, Genomic_DNA.
 DR SMR: Q6JC10; 217-736.
 DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 736 AA; 81388 MW; BA9463E970028DFO CRC64;

Query Match 96.6%; Score 3853; DB 2; Length 736;
 Best Local Similarity 96.6%; Pred. No. 7.3e-241;
 Matches 711; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

OY	1	MAAGCYLPDMLENDLSGIRREWDLKPGAPKPKANQKODDORGLVLPGYKYLGPFGND	60
DB	1	MAAGCYLPDMLENDLSGIRREWDLKPGAPKPKANQKODDORGLVLPGYKYLGPFGND	60
OY	61	KGEFVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGGNGRAVFO	120
DB	61	KGEFVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGGNGRAVFO	120
OY	121	AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSGIGTKGQOPAKKRLNFGQTGDS	180
DB	121	AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQEPDSSGIGTKGQOPAKKRLNFGQTGDS	180
OY	181	SVPDPQPLGEPAPATPAAGPTTMAAGGAPMADNNEGADGVNAGNHCDSITWLGDRVI	240
DB	181	SVPDPQPLGEPAPATPAAGPTTMAAGGAPMADNNEGADGVNAGNHCDSITWLGDRVI	240
OY	241	TTSTRTMAALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDNRRHCHFSPPDMQRL	300
DB	241	TTSTRTMAALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDNRRHCHFSPPDMQRL	300
OY	301	INNMGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQ	360
DB	301	INNMGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQ	360
OY	361	GCLPPFPADVPMIPOYGITLTNNGSOAVGRSSFCLEFPQMLRTGNFTFSYFEEVP	420
DB	361	GCLPPFPADVPMIPOYGITLTNNGSOAVGRSSFCLEFPQMLRTGNFTFSYFEEVP	420
OY	421	FHSSVAHSOSLDRLNMPILIDQYLYLTNRTONQSGAONKDLFSRGSPPAGMSVOPKMLP	480
DB	421	FHSSVAHSOSLDRLNMPILIDQYLYLTNRTONQSGAONKDLFSRGSPPAGMSVOPKMLP	480
OY	481	GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMGCV	540
DB	481	GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMGCV	540
OY	541	MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPATGTVHAMG	600
DB	541	MIFGKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVANFOSSSTDPATGTVHAMG	600

```
QY 601 ALPGVWQDRDVLVYLCGPIMAKI PHTDGHFHPSPLMGFGKLNPPPOILIKNTVPANPPA 660
DB 601 ALPGVWQDRDVLVYLCGPIMAKI PHTDGHFHPSPLMGFGKLNPPPOILIKNTVPANPPA 660
QY 661 EESATKFAFSTIQTGVSQVSEIEMELQKENS KRNWPEVQTSNYSKANSANVPTVDNNG 720
DB 661 EESATKFAFSTIQTGVSQVSEIEMELQKENS KRNWPEVQTSNYSKANSANVPTVDNNG 720
QY 721 YTEPRPIGTRYLTRPL 736
DB 721 YTEPRPIGTRYLTRPL 736

RESULT 7
056139_9VIRU PRELIMINARY; PRT; 736 AA.
ID 056139_9VIRU PRELIMINARY; PRT; 736 AA.
AC 056139_
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Capsid protein VPI.
OS Adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_Taxid=68742;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98080418; PubMed=9420229;
RUTledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
RT (AAV) serotypes other than AAV type 2."
RT J. Virol. 72:309-319 (1998).
DR EMBL; AF028705; AAB95452.1; -; Genomic_DNA.
DR SMR; 056139; 217-736.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81906 MW; DD52331AD5FD070F CRC64;

Query Match 88.0%; Score 3511; DB 2; Length 736;
Best Local Similarity 86.8%; Pred. No. 1e-216;
Matches 640; Conservative 38; Mismatches 57; Indels 2; Gaps 2;
```

```
DB 420 FHSSVAHSQSLDRMLNPLIDQVLYLNRTO- NQSGAQNOKDLFSRGS PAKMSVQAPKWL 479
QY 480 PGPCTRGQRPVSKTCTDNNNSNFTGASKNLNGEESLIINPTAAASHKDDDKFFPMG 539
DB 480 PGPCTRGQRLSTXANDNNNSNFPMTAAASKYHLNGRDSLVNPGPAAASHKDDDKFFPMG 539
QY 540 VMIPEKESAGSNTLMDVMTDEEIKATNPVATERFGTVAVNFOSSSTDPATGDVHAM 599
DB 540 NLIIFKEGCTTASNAELDVMITDEEIRTNNEVATEIGTVANNLOSSNTAFTTTVDNQ 599
QY 600 GALPGVWQDRDVLVYLCGPIMAKI PHTDGHFHPSPLMGFGKLNPPPOILIKNTVPANPP 659
DB 600 GALPGVWQDRDVLVYLCGPIMAKI PHTDGHFHPSPLMGFGKLNPPPOILIKNTVPANPP 659
QY 660 AEFSAKFAFSTIQTGVSQVSEIEMELQKENS KRNWPEVQTSNYSKANSANVPTVDNNG 719
DB 660 TTFSPAKFAFSTIQTGVSQVSEIEMELQKENS KRNWPEIQTYSNYSKANSANVPTVDNNG 719
QY 720 LYTEPRPIGTRYLTRPL 736
DB 720 YTEPRPIGTRYLTRPL 736

RESULT 8
065311_9VIRU PRELIMINARY; PRT; 736 AA.
ID 065311_9VIRU PRELIMINARY; PRT; 736 AA.
AC 065311_
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 3.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_Taxid=46350;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96266430; PubMed=8661429; DOI=10.1006/viro.1996.0367;
RUMuramatsu S., Mizukami H., Young N.S., Brown K.B.;
RT "Nucleotide sequencing and generation of an infectious clone of adeno-
RT associated virus 3."
RT Virology 221:208-217 (1996).
DR EMBL; U48704; AAC55049.1; -; Genomic_DNA.
DR SMR; 065311; 217-736.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81660 MW; APTLEF47B5C67A10 CRC64;

Query Match 87.6%; Score 3494; DB 2; Length 736;
Best Local Similarity 86.4%; Pred. No. 1.3e-217;
Matches 637; Conservative 39; Mismatches 59; Indels 2; Gaps 2;
```

```
OY 301 INNMGFRPKRLNFKLFIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPVYLGSANQ 360
DB 300 INNMGFRPKLSFLFLFIQVKGVTQNDGTTIANNLSTVQVFTDSYQLPYVLGSANQ 359
OY 361 GCLPPPADVFMVPOYGYLTLLNGSQA VGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 420
DB 360 GCLPPPADVFMVPOYGYLTLLNGSQA VGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 419
OY 421 FHSSVYAHQSOLDRLNPLIDQYLYLNRTOGTTSGTTSQSRLFSQAQPSQSLQARWML 479
DB 420 FHSSVYAHQSOLDRLNPLIDQYLYLNRTOGTTSGTTSQSRLFSQAQPSQSLQARWML 479
OY 480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPM5G 539
DB 480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPM5G 539
OY 540 VMI FCKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 599
DB 540 VMI FCKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 599
OY 600 GALPGMWQODRDVYLOGPIMAKIPHTDGHFHSPLMGFGKXNPPOLLIKXTVPVAPNP 659
DB 600 GALPGMWQODRDVYLOGPIMAKIPHTDGHFHSPLMGFGKXNPPOLLIKXTVPVAPNP 659
OY 660 AEFSAITKFASTITQYSTQGVSEIEMELOKENSXKRNPEVOYTSNYAKSANVDFTVDNNG 719
DB 660 TTFSAKAPASFTITQYSTQGVSEIEMELOKENSXKRNPEIQTYSNYSKSNVAVDFTVDNNG 719
OY 720 LYTEPRPIGTRYLTRPL 736
DB 720 VYSEBRPIGTRYLTRNL 736

RESULT 9
OY Q608Y3_9VIRU PRELIMINARY; PRT; 736 AA.
AC Q608Y3;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Capid protein.
GN Name=VPI;
OS Non-human primate Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
CX NCBI_TaxID=226582;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22631657; PubMed=12716974; DOI=10.1073/pnas.0937739100;
RA Gao G., Alvirra M.R., Samanathan S., Lu Y., Vandenbergh L.H.,
RT "Adeno-associated viruses undergo substantial evolution in primates
RT during natural infections."
RT Proc. Natl. Acad. Sci. U.S.A. 100:6081-6086(2003).
DR EMBL; AY242997; AA08183.1; -; Genomic_DNA.
DR SMR; Q608Y3; 217-736.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat.1.
SQ SEQUENCE 736 AA; 81370 MW; 87D05047F8FA14C CRC64;

Query Match 87.1%; Score 3473; DB 2; Length 736;
Best Local Similarity 85.6%; Pred. No. 3e-216;
Matches 631; Conservative 44; Mismatches 60; Indels 2; Gaps 2;
```

```
OY 121 AKKVLBLEGLVEGAKTAPGKRPVQSPQEPSSSGIGKTGQAPAKKLNFGQTDSE 180
DB 121 AKKVLBLEGLVEGAKTAPGKRPVQSPQEPSSSGIGKTGQAPAKKLNFGQTDSE 180
OY 181 SVDPQPLGEBPATPAVGFPTMASGGQAPMADNBSAGDVGNAGWNCDSITMLGDRVI 240
DB 181 SVDPQPLGEBPATPAVGFPTMASGGQAPMADNBSAGDVGNAGWNCDSITMLGDRVI 240
OY 241 TTSRTMALPTYNHNLKQISSASTGAS -NDNHFGYSTPMGVDFNRFCHSPRDMOR 299
DB 241 TTSRTMALPTYNHNLKQISSASTGAS -NDNHFGYSTPMGVDFNRFCHSPRDMOR 299
OY 300 LINNMGFRPKRLNFKLFIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPVYLGSAN 359
DB 301 LINNMGFRPKRLNFKLFIQVKEVTTNEGTIANNLSTVQVFTDSYQLPYVLGSAN 360
OY 360 GCLPPPADVFMVPOYGYLTLLNGSQA VGRSSFCLEYPFSQMLRTGNNFTFSYTFEEV 419
DB 361 GCLPPPADVFMVPOYGYLTLLNGSQA VGRSSFCLEYPFSQMLRTGNNFTFSYTFEEV 420
OY 420 FHSSVYAHQSOLDRLNPLIDQYLYLNRTOGTTSGTTSQSRLFSQAQPSQSLQARWML 479
DB 421 FHSSVYAHQSOLDRLNPLIDQYLYLNRTOGTTSGTTSQSRLFSQAQPSQSLQARWML 479
OY 480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPM5G 539
DB 480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPM5G 539
OY 540 VMI FCKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 599
DB 540 VMI FCKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 599
OY 600 GALPGMWQODRDVYLOGPIMAKIPHTDGHFHSPLMGFGKXNPPOLLIKXTVPVAPNP 659
DB 600 GALPGMWQODRDVYLOGPIMAKIPHTDGHFHSPLMGFGKXNPPOLLIKXTVPVAPNP 659
OY 660 AEFSAITKFASTITQYSTQGVSEIEMELOKENSXKRNPEVOYTSNYAKSANVDFTVDNNG 719
DB 660 LTFNQAUKNSFTITQYSTQGVSEIEMELOKENSXKRNPEIQTYSNYSKSNVAVDFTVDNNG 719
OY 720 LYTEPRPIGTRYLTRPL 736
DB 720 VYSEBRPIGTRYLTRNL 736

RESULT 10
OY Q670O8_9VIRU PRELIMINARY; PRT; 735 AA.
AC Q670O8;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Capid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
CX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chen C.L., Jensen R.L., Schnepf B.C., Connell M.J., Bartlett J.S.,
RT "Characterization of Adeno-Associated Viruses in Children."
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY695376; AAU05370.1; -; Genomic_DNA.
DR SMR; Q670O8; 217-735.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat.1.
SQ SEQUENCE 735 AA; 81867 MW; 258BF955F1C9925 CRC64;

Query Match 87.0%; Score 3469.5; DB 2; Length 735;
```


Best Local Similarity 85.1%; Pred. No. 56-216; Matches 626; Conservative 47; Mismatches 62; Indels 1; Gaps 1;

QY 1 MAADGYPDWLBDNISGIRWMDLKPAPKPKANQKODDGRGLVPGYKYLGFPGNGD 60
DB 1 MAADGYPDWLBDNISGIRWMDLKPAPKPKAPRHRHDDRGVLPGYKYLGFPGNGD 60
QY 61 KCEPVNAADAAALHDKAYDQOLKAGDNPYLRYNHADADEFORLLEDTSFGGNLGRAVQ 120
DB 61 KCEPVNAADAAALHDKAYDQOLKAGDNPYLRYNHADADEFORLLEDTSFGGNLGRAVQ 120
QY 121 AKKRVLEPGLVEEGAKTAPGKKRPVEOSPOBDSGGIGKTGQOQPAKKRLNFGQTGSE 180
DB 121 AKKRVLEPGLVEEGAKTAPGKKRPVEHSPBDSGGIGKTGQOQPARRLNFGQTGSD 180
QY 181 SVDPDQPLGEPAPATPAVGPPTTMASSGGA PMADNNEGADGVNAGSNHMCDSQWLGDVY 240
DB 181 SVDPDQPLGEPAPATPAVGPPTTMASSGGA PMADNNEGADGVNAGSNHMCDSQWLGDVY 240
QY 241 TTSTRTVALPTYNNHLYKQISSASTGASNDNHFGYSTPMGYFDENRFCHFSPRDWQRL 300
DB 241 TTSTRTVALPTYNNHLYKQISSQS -GASNDNHFGYSTPMGYFDENRFCHFSPRDWQRL 299
QY 301 INNNNGEPRKRLNFKLFPNIQVKEVTNDGVTIANNLSTGYVFSSEYQLPYVLSAQ 360
DB 301 INNNNGEPRKRLNFKLFPNIQVKEVTNDGVTIANNLSTGYVFTDSEYQLPYVLSAQ 359
QY 361 GCLPPPADVPMIPOGYLTLLNNGSOAVGRSSFCLEYPQOMLRTGNNFTFSYTFEEVP 420
DB 361 GCLPPPADVPMIPOGYLTLLNNGSOAVGRSSFCLEYPQOMLRTGNNFTFSYTFEDVP 419
QY 421 FHSSYAHQSGLRLMNPILIDQYLYLNRTQONSGAQNKDLPFSRGS PAMGVSQPKWMLP 480
DB 421 FHSSYAHQSGLRLMNPILIDQYLYLNRTQONSGAQTQOGRILFSGAGPTMSLQAQKWL 479
QY 481 GPCYRQORVSKTKTNNNSNFTTGTGASKYNLGRRESIINPGTAMASHKXDEDEKFPMSGV 540
DB 481 GPCYRQORVSKTKTNNNSNFTTGTGASKYNLGRRESIINPGTAMASHKXDEDEKFPMSGV 539
QY 541 MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHANG 600
DB 541 MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHANG 599
QY 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKPNPPIILIKNTPVPANPPA 660
DB 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKPNPPIILIKNTPVPANPPA 659
QY 661 EFSATKFAFTIQTSGQVSEIEMWELQKENSKRANPEVOYTSNYAKSANVDTVDNGL 720
DB 661 EFSATKFAFTIQTSGQVSEIEMWELQKENSKRANPEIQTSTNYKSVNVDFTVDNGLV 719
QY 721 YTEPRPIGTRYLTRPL 736
DB 721 YTEPRPIGTRYLTRNL 735

RESULT 11
Q6JB27_9VIRU
ID Q6JB27_9VIRU PRELIMINARY; PRT: 735 AA.
AC Q6JB27;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=cap;
OC Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN NUCLEOTIDE SEQUENCE.
RP PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X., Wilson J.M.;

RT "Clades of Adeno-associated viruses are widely disseminated in human tissues".
RT U. Virol. 78:6381-6388(2004).
DR EMBL: AY530622; AAS99307.1; -; Genomic_DNA.
DR SMR: Q6JB27; 217-735.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parv_coat.
DR Pfam: PF00740; Parv_coat; 1.
SQ SEQUENCE 735 AA; 81947 MW; EBA318E904E0BF0A CRC64;

Query Match 86.7%; Score 3456.5; DB 2; Length 735;
Best Local Similarity 84.5%; Pred. No. 3-56-215;
Matches 622; Conservative 49; Mismatches 64; Indels 1; Gaps 1;

QY 1 MAADGYPDWLBDNISGIRWMDLKPAPKPKANQKODDGRGLVPGYKYLGFPGNGD 60
DB 1 MAADGYPDWLBDNISGIRWMDLKPAPKPKAPRHRHDDRGVLPGYKYLGFPGNGD 60
QY 61 KCEPVNAADAAALHDKAYDQOLKAGDNPYLRYNHADADEFORLLEDTSFGGNLGRAVQ 120
DB 61 KCEPVNAADAAALHDKAYDQOLKAGDNPYLRYNHADADEFORLLEDTSFGGNLGRAVQ 120
QY 121 AKKRVLEPGLVEEGAKTAPGKKRPVEOSPOBDSGGIGKTGQOQPAKKRLNFGQTGSE 180
DB 121 AKKRVLEPGLVEEGAKTAPGKKRPVEHSPBDSGGIGKTGQOQPARRLNFGQTGSD 180
QY 181 SVDPDQPLGEPAPATPAVGPPTTMASSGGA PMADNNEGADGVNAGSNHMCDSQWLGDVY 240
DB 181 SVDPDQPLGEPAPATPAVGPPTTMASSGGA PMADNNEGADGVNAGSNHMCDSQWLGDVY 240
QY 241 TTSTRTVALPTYNNHLYKQISSASTGASNDNHFGYSTPMGYFDENRFCHFSPRDWQRL 300
DB 241 TTSTRTVALPTYNNHLYKQISSQS -GASNDNHFGYSTPMGYFDENRFCHFSPRDWQRL 299
QY 301 INNNNGEPRKRLNFKLFPNIQVKEVTNDGVTIANNLSTGYVFSSEYQLPYVLSAQ 360
DB 301 INNNNGEPRKRLNFKLFPNIQVKEVTNDGVTIANNLSTGYVFTDSEYQLPYVLSAQ 359
QY 361 GCLPPPADVPMIPOGYLTLLNNGSOAVGRSSFCLEYPQOMLRTGNNFTFSYTFEEVP 420
DB 361 GCLPPPADVPMIPOGYLTLLNNGSOAVGRSSFCLEYPQOMLRTGNNFTFSYTFEDVP 419
QY 421 FHSSYAHQSGLRLMNPILIDQYLYLNRTQONSGAQNKDLPFSRGS PAMGVSQPKWMLP 480
DB 421 FHSSYAHQSGLRLMNPILIDQYLYLNRTQONSGAQTQOGRILFSGAGPTMSLQAQKWL 479
QY 481 GPCYRQORVSKTKTNNNSNFTTGTGASKYNLGRRESIINPGTAMASHKXDEDEKFPMSGV 540
DB 481 GPCYRQORVSKTKTNNNSNFTTGTGASKYNLGRRESIINPGTAMASHKXDEDEKFPMSGV 539
QY 541 MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHANG 600
DB 541 MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHANG 599
QY 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKPNPPIILIKNTPVPANPPA 660
DB 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKPNPPIILIKNTPVPANPPA 659
QY 661 EFSATKFAFTIQTSGQVSEIEMWELQKENSKRANPEVOYTSNYAKSANVDTVDNGL 720
DB 661 EFSATKFAFTIQTSGQVSEIEMWELQKENSKRANPEIQTSTNYKSVNVDFTVDNGLV 719
QY 721 YTEPRPIGTRYLTRPL 736
DB 721 YTEPRPIGTRYLTRNL 735

RESULT 12
Q6JCS2_9VIRU
ID Q6JCS2_9VIRU PRELIMINARY; PRT: 737 AA.
AC Q6JCS2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxId=272636;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
Wilson J.M.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues."
RT J. Virol. 78:6381-6388(2004).
DR EMBL: AY530567; AAS99252.1; -; Genomic_DNA.
DR SMR: Q6JC52; 218-737.
DR GO: GO:0019028; C:viral capsid, IEA.
DR GO: GO:0005198; F:structural molecule activity, IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat/1.
SQ SEQUENCE 737 AA; 81538 MW; CC68B97BE1F25C CRC64;

Query Match 86.7%; Score 3456.5; DB 2; Length 737;
Best Local Similarity 85.8%; Pred. No. 3.5e-215;
Matches 634; Conservative 47; Mismatches 53; Indels 5; Gaps 4;

QY 1 MAADGYLDPMLENDLSBSGIREMWDLPKGPAPKXNQQKODDGRGLVPGYKYLGPFGNLD 60
DB 1 MAADGYLDPMLENDLSBSGIREMWDLPKGPAPKXNQQKODDGRGLVPGYKYLGPFGNLD 60
QY 61 KGEFVNADAAALHDKAYDQOLKAGDNPLYRYNHADAEPQERLOEDTSFGGNLGRAVFO 120
DB 61 KGEFVNADAAALHDKAYDQOLKAGDNPLYRYNHADAEPQERLOEDTSFGGNLGRAVFO 120
QY 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQ-EPDSSSGICTGQQQPAKKRLNFGQTGDS 179
DB 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQ-EPDSSSGICTGQQQPAKKRLNFGQTGDS 180
QY 180 ESVDPQPIGEPBPAPPAVGPPTMASGGGAPMADNBEAGDVGNAAGMWHCDSTWLGDRV 239
DB 180 ESVDPQPIGEPBPAPPAVGPPTMASGGGAPMADNBEAGDVGNAAGMWHCDSTWLGDRV 240
QY 240 ITTSTRTALPTYNNHLKYKQISSASTGASNDNHYFGYSTPMGYDPFNHFGHFSFRDQR 299
DB 240 ITTSTRTALPTYNNHLKYKQISSASTGASNDNHYFGYSTPMGYDPFNHFGHFSFRDQR 300
QY 300 LINNMGFRPKLNFKLFNIOVKEVTTNDGVTIANNLITSTVOYFSDSEYOLPYVLGSAH 359
DB 300 LINNMGFRPKLNFKLFNIOVKEVTTNDGVTIANNLITSTVOYFSDSEYOLPYVLGSAH 360
QY 360 QGCLPPFPADVPMIPOYGYLTNNNGSOGVSGSFCLFEPFSOMLRTGNNTFSSTFEV 419
DB 360 QGCLPPFPADVPMIPOYGYLTNNNGSOGVSGSFCLFEPFSOMLRTGNNTFSSTFEV 420
QY 420 PPHSSYASOSLDRLNPLIDQYLYLARTQ-NOSGSAQNKDLSRSGSPAGMSYQPKW 478
DB 420 PPHSSYASOSLDRLNPLIDQYLYLARTQ-NOSGSAQNKDLSRSGSPAGMSYQPKW 480
QY 479 LFGPCYRQORVSKTLDNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMS 538
DB 479 LFGPCYRQORVSKTLDNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMS 540
QY 539 GVMIGKESAGASN-TALDNWMTDBEETKATNPVATERFGTVANFQSSSTDPATGVH 597
DB 539 GVMIGKESAGASN-TALDNWMTDBEETKATNPVATERFGTVANFQSSSTDPATGVH 598
QY 541 GVLITGK-TGALNKTLENTLMEETLRPTNPVATEBYGLVSSNLQAAAOOTGVN 598
DB 541 GVLITGK-TGALNKTLENTLMEETLRPTNPVATEBYGLVSSNLQAAAOOTGVN 598
QY 598 AMGALPGMWQODRDVYLGPIWAKIPHTDGHFSPMLMGFGKPKPPDILIKNTPVAN 657
DB 598 AMGALPGMWQODRDVYLGPIWAKIPHTDGHFSPMLMGFGKPKPPDILIKNTPVAN 658
QY 658 PPAEFSATKFAFTOYSTGVSVBIEMELQKENSRRNPEVOYTSNAXKSNANDFTVDN 717
DB 658 PPAEFSATKFAFTOYSTGVSVBIEMELQKENSRRNPEVOYTSNAXKSNANDFTVDN 717

DB 659 PPEVFPKAFASFTTOYSTGVSVBIEMELQKENSRRNPEIYTSNFDKOTGVDAVDS 718
QY 718 NGLYTEPPPICTRYLTRPL 736
DB 719 QGVYSERPPICTRYLTRNL 737

RESULT 13
Q6JC58_VIRU PRELIMINARY; PRT; 737 AA.
ID Q6JC58_VIRU
AC Q6JC58
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxId=272636;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
Wilson J.M.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues."
RT J. Virol. 78:6381-6388(2004).
DR EMBL: AY530561; AAS99246.1; -; Genomic_DNA.
DR SMR: Q6JC58; 218-737.
DR GO: GO:0019028; C:viral capsid, IEA.
DR GO: GO:0005198; F:structural molecule activity, IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat/1.
SQ SEQUENCE 737 AA; 81486 MW; EF6B85A99F07CB5 CRC64;

Query Match 86.6%; Score 3455.5; DB 2; Length 737;
Best Local Similarity 85.6%; Pred. No. 4.1e-215;
Matches 632; Conservative 48; Mismatches 55; Indels 3; Gaps 3;

QY 1 MAADGYLDPMLENDLSBSGIREMWDLPKGPAPKXNQQKODDGRGLVPGYKYLGPFGNLD 60
DB 1 MAADGYLDPMLENDLSBSGIREMWDLPKGPAPKXNQQKODDGRGLVPGYKYLGPFGNLD 60
QY 61 KGEFVNADAAALHDKAYDQOLKAGDNPLYRYNHADAEPQERLOEDTSFGGNLGRAVFO 120
DB 61 KGEFVNADAAALHDKAYDQOLKAGDNPLYRYNHADAEPQERLOEDTSFGGNLGRAVFO 120
QY 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQ-EPDSSSGICTGQQQPAKKRLNFGQTGDS 179
DB 121 AKKRVLEPLGLVEBGAKTAPGKKRPVEQSPQ-EPDSSSGICTGQQQPAKKRLNFGQTGDS 180
QY 180 ESVDPQPIGEPBPAPPAVGPPTMASGGGAPMADNBEAGDVGNAAGMWHCDSTWLGDRV 239
DB 180 ESVDPQPIGEPBPAPPAVGPPTMASGGGAPMADNBEAGDVGNAAGMWHCDSTWLGDRV 240
QY 240 ITTSTRTALPTYNNHLKYKQISSASTGASNDNHYFGYSTPMGYDPFNHFGHFSFRDQR 299
DB 240 ITTSTRTALPTYNNHLKYKQISSASTGASNDNHYFGYSTPMGYDPFNHFGHFSFRDQR 300
QY 241 ITTSTRTALPTYNNHLKYKQISSASTGASNDNHYFGYSTPMGYDPFNHFGHFSFRDQR 300
DB 241 ITTSTRTALPTYNNHLKYKQISSASTGASNDNHYFGYSTPMGYDPFNHFGHFSFRDQR 300
QY 300 LINNMGFRPKLNFKLFNIOVKEVTTNDGVTIANNLITSTVOYFSDSEYOLPYVLGSAH 359
DB 300 LINNMGFRPKLNFKLFNIOVKEVTTNDGVTIANNLITSTVOYFSDSEYOLPYVLGSAH 360
QY 360 QGCLPPFPADVPMIPOYGYLTNNNGSOGVSGSFCLFEPFSOMLRTGNNTFSSTFEV 419
DB 360 QGCLPPFPADVPMIPOYGYLTNNNGSOGVSGSFCLFEPFSOMLRTGNNTFSSTFEV 420
QY 420 PPHSSYASOSLDRLNPLIDQYLYLARTQ-NOSGSAQNKDLSRSGSPAGMSYQPKW 478
DB 420 PPHSSYASOSLDRLNPLIDQYLYLARTQ-NOSGSAQNKDLSRSGSPAGMSYQPKW 480
QY 479 LFGPCYRQORVSKTLDNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMS 538
DB 479 LFGPCYRQORVSKTLDNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMS 538


```
Db 481 LPGPCRQQRVSCTLDQNNNSNFAMTGA TKHILNGRSLVNPGVAMATHKDDEERFPSS 540
Qy 539 GVMIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVA VNFQSSSTDPATGDVHA 598
Db 541 GVLIFGKTGA -ANKTTLENVLMNTNEBEIRPTNPVATEEYGT VSNLQANATAAQOVVNN 599
Qy 599 MGLALGMMWQDDVYLOGPIWAKIPHTDGHFHSPLMGFGFLK NPPOLLIKNTVPANP 658
Db 600 QGALPGMMWQNRDYYLOGPIWAKIPHTDGNFHSPLMGFGFLK HPPOLLIKNTVPANP 659
Qy 659 PAFSATKFASTFOYSTGQVSVEIEMELQKENSKRNNPEVOYT SNTVAKSANVDFTVDN 718
Db 660 PEVFPFAKFASTFOYSTGQVSVEIEMELQKENSKRNNPEI OYTSNFKQGTVDRAVDSQ 719
Qy 719 GLTEPRPIGTRVLTREPL 736
Db 720 GYSEPRPIGTRVLTREPL 737

RESULT 14
06JC28_9VIRU PRELIMINARY; PRT; 735 AA.
ID 06JC28_9VIRU PRELIMINARY; PRT; 735 AA.
AC 06JC28_9VIRU PRELIMINARY; PRT; 735 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_Taxid=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530591; AAS99276.1; -; Genomic_DNA.
DR SMR; Q6JC28; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81897 MW; 197066F8911FAE9B CRC64;

Query Match 86.6%; Score 3453.5; DB 2; Length 735;
Best Local Similarity 84.5%; Pred. No. 5.5e-215;
Matches 622; Conservative 48; Mismatches 65; Indels 1; Gaps 1;
```

```
Db 300 INNMGFRPKRLNFKLNLIVKVEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 359
Qy 361 GCLPFPADVMIPIQYGLTLTNKSGQAVGRSSFYLETFFPSQMLATGNNTFFSYFEEDVP 420
Db 360 GCLPFPADVMIPIQYGLTLTNKSGQAVGRSSFYLETFFPSQMLATGNNTFFSYFEEDVP 419
Qy 421 FHSYAHQSQSLDRLMNPILDOYLYLNTQNOGSAOKDILLFSRGS PAGMSVCPKMLP 480
Db 420 FHSYAHQSQSLDRLMNPILDOYLYLNTQNOGSAOKDILLFSRGS PAGMSVCPKMLP 479
Qy 481 GPCYRQQRVSCTLDQNNNSNFTMGASXYNLNGRESINPGTMA SHKDEDEKFPMSGV 540
Db 480 GPCYRQQRVSCTLDQNNNSNFTMGASXYNLNGRESINPGTMA SHKDEDEKFPMSGV 539
Qy 541 MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVA VNFQSSSTDPATGDVHA 600
Db 540 LIFGKTGANADADLENWMTDEBEIRPTNPVATEEYGT VSNLQANATAAQOVVNN 599
Qy 601 ALPGMMWQDDVYLOGPIWAKIPHTDGHFHSPLMGFGFLK NPPOLLIKNTVPANP 660
Db 600 ALPGMMWQDDVYLOGPIWAKIPHTDGHFHSPLMGFGFLK NPPOLLIKNTVPANP 659
Qy 661 EFSATKFASTFOYSTGQVSVEIEMELQKENSKRNNPEVOYT SNTVAKSANVDFTVDN 720
Db 660 NFSAKFASTFOYSTGQVSVEIEMELQKENSKRNNPEI OYTSNFKQGTVDRAVDSQ 719
Qy 721 YTEPRPIGTRVLTREPL 736
Db 720 YSEPRPIGTRVLTREPL 735

RESULT 15
06JC19_9VIRU PRELIMINARY; PRT; 738 AA.
ID 06JC19_9VIRU PRELIMINARY; PRT; 738 AA.
AC 06JC19_9VIRU PRELIMINARY; PRT; 738 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_Taxid=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530600; AAS99285.1; -; Genomic_DNA.
DR SMR; Q6JC19; 218-738.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 738 AA; 81582 MW; 585F46CF0769934F CRC64;

Query Match 86.6%; Score 3453; DB 2; Length 738;
Best Local Similarity 85.0%; Pred. No. 5.9e-215;
Matches 627; Conservative 48; Mismatches 61; Indels 2; Gaps 2;
```

```

Db      121 AKRVLBEPLGLVEBAKTA PGKRPVEBPS PORSPDS TGIGKGOQPAKKRLNFGQTGDS 180
QY      180 ESVDPOLGEPBPATPAVPTTMA SGGAPMADNNEGADGVGNASGNMHCSTWLGDRV 239
Db      181 ESVDPOLGEPBPAGSGTMA GAGAPMADNNEGADGVGSSGNMHCSTWLGDRV 240
QY      240 ITTSTRTVALPTYNHLYKQIS SASTGAS -NDNHFGYSTPWGYPDFNRFCHFSPRDMQ 298
Db      241 ITTSTRTVALPTYNHLYKQIS NGTSGSSTNDNTYFGYSTPWGYDFNRFCHFSPRDMQ 300
QY      299 RLINNNMGRRPRRLNFKLFNI QVKEYTNDGYTTIANNLSTVQYFSDSEYQLPYVLGSA 358
Db      301 RLINNNMGRRPRRLNFKLFNI QVKEYTNDGYTTIANNLSTVQYFSDSEYQLPYVLGSA 360
QY      359 HOGCLPRPADVFMIPQGYLT LNNGSOAVGRSSFYCLEYFPSPQMLRTGNNPFSYTEEB 418
Db      361 HOGCLPRPADVFMIPQGYLT LNNGSOAVGRSSFYCLEYFPSPQMLRTGNNPFSYTEEB 420
QY      419 VPFHSSYAHSGSLDRLMNP LIQVLYLNRTOGSGSAQNKDLFSGSPAGMSVQPRNM 478
Db      421 VPFHSSYAHSGSLDRLMNP LIQVLYLNRTOGSGTGTQQLFSGAGPAMMSAQAKNM 480
QY      479 LFGPCYRQQRVSKTKTDNN NSNFTWTGASKYNLNGRESTINFGTAMASHKDEDEKFFPMS 538
Db      481 LFGPCYRQQRVSKTKTDNN NSNFTWTGASKYNLNGRESTINFGTAMASHKDEDEKFFPMS 540
QY      539 GVMIFGKESAGASNTALDNV MTDEEIKATNPVATERFGTVAVNFQSSSTDPAQDVHA 598
Db      541 GVMIFGKQAGARDNVYSVM LTBSEELKTNPVATEBOYGVVADNLQOTNTGPIVGNVS 600
QY      599 MGALFGMWQDRDVLQGPIM AKIPHTDGHFHSPLMGFGGLKNPPOILLKNTVPANP 658
Db      601 OGALFGMWQDRDVLQGPIM AKIPHTDGNFHSPLMGFGGLKNPPOILLKNTVPANP 660
QY      659 PAEFSATKFASTITQYSTG QVSEIEMELQKENS KRMNPEVOYTSNVAKSANVDFTVDNN 718
Db      661 PTFFSQAKLASFITQYSTG QVSEIEMELQKENS KRMNPEIQTYSNYYKSTVNDFAVNT 720
QY      719 GLYTEPRPIGTRYLRPL 736
Db      721 GTYSEPRPIGTRYLRNL 738

```

Search completed: November 23, 2005, 17:41:16
 Job time : 144.766 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:29:19 ; Search time 25.9904 Seconds
(without alignments)
2724.684 Million cell updates/sec

Title: US-10-696-282-13

Sequence: 3989 1 MAADGYLPDWMEDNLSEGR.....NNGLYTEPRPIGTRVLRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	2177.5	54.6	504	1 VCPV3A	coat protein - ade
2	2168.5	54.4	732	2 S52210	coat protein VP1 -
3	798.5	20.0	673	1 VCPV3B	coat protein VP1 -
4	600.5	15.1	781	1 VCPV19	coat protein VP1 -
5	482	12.1	729	1 A60006	coat protein VP1 -
6	482	12.1	729	1 VCPVNA	coat protein VP1 -
7	444.5	11.0	727	1 VCPV1F	coat protein VP1 -
8	439.5	11.0	727	1 VCPVME	coat protein VP1 -
9	431.5	10.8	727	1 VCPVFP	coat protein VP1 -
10	426	10.7	713	1 VCPVPP	coat protein VP1 -
11	424.5	10.6	718	1 VCPVIM	coat protein VP1 -
12	421.5	10.6	748	1 VCPVCP	coat protein VP1 -
13	407.5	10.2	737	1 VCPVCD	coat protein VP1 -
14	404	10.1	722	1 VCPVY2	coat protein VP1 -
15	385.5	9.7	722	1 VCPV2M	coat protein VP1 -
16	377	9.5	716	1 VCPV2M	coat protein VP1 -
17	234	5.9	587	1 B44276	coat protein VP1 -
18	199.5	5.0	810	2 A44054	orf1 protein - Jun
19	198	5.0	584	2 S49594	capsid protein VP2
20	191	4.8	702	1 VCPVAP	coat protein VP1 -
21	143.5	3.6	648	2 S50856	whn protein - rat
22	139	3.5	931	2 T49710	related to glucan
23	131.5	3.3	1446	1 A45344	immediate-early pr
24	130.5	3.3	1145	2 T18235	transcription acti
25	130	3.3	1338	2 T30565	MAP kinase kinase
26	129	3.2	1072	2 A86827	hypothetical prote
27	124.5	3.1	2493	2 A55481	adenylate cyclase
28	124	3.1	1296	2 C82521	hemolysin-type cal
29	124	3.1	3078	2 T28432	variant-specific s

30	123.5	3.1	690	2 AB0124	probable TonB-depe
31	123.5	3.1	2529	2 B64635	toxin-like outer m
32	123	3.1	1142	2 T37435	enamelin precursor
33	122.5	3.1	880	1 SYBSVS	valine-tRNA ligase
34	121	3.0	635	2 P9660	protein F2K11.10 f
35	121	3.0	1113	2 S28925	nuclear pore compl
36	121	3.0	1549	2 T48103	type VII collagen
37	120	3.0	1379	2 T37752	hypothetical serin
38	119.5	3.0	1180	2 B86719	hypothetical prote
39	119.5	3.0	1394	2 S60762	IgA-specific serin
40	119.5	3.0	2004	2 AC0314	probable membrane
41	119	3.0	667	2 A41311	transcription fact
42	118.5	3.0	2399	2 H71879	toxin-like outer m
43	118	3.0	739	2 T52289	probable transkeo
44	118	3.0	956	2 T08144	myrosinase-binding
45	117.5	2.9	1076	1 A35622	nuclear pore prote

ALIGNMENTS

RESULT 1

coat protein - adeno-associated virus type 2

C/Species: adeno-associated virus type 2

C/Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004

C/Accession: A03698

R/Srivastava, A.; Lusby, E.W.; Berns, K.I.

J. Virol. 45, 555-564, 1983

A/Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.

A/Reference number: A03694; MUID:83164299; PMID:6300419

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-504 <SRI>

A/Cross-references: UNIPROT:P03135; UNIPARC:UPI0000127C9E; EMBL:J01901; NID:9209616; PI

C/Superfamily: adeno-associated virus coat protein

C/Keywords: coat protein

Query Match 54.6%; Score 2177.5; DB 1; Length 504;

Best Local Similarity 82.5%; Pred. No. 3.9e-134;

Matches 392; Conservative 33; Mismatches 47; Indels 3; Gaps 2;

QY	203	MASGGAPMADNNEGADGVGASGWHCDSTWLGDRVITTSRTMALPTYNHLYQISS	262
DB	1	MATGSGAPMADNNEGADGVGASGWHCDSTWLGDRVITTSRTMALPTYNHLYQISS	60
QY	263	ASTGASNDNHYFGYSTPMGYPDPNRFCHFSRDMQRLINNNMGFPKRLNPLFIQYK	322
DB	61	QS-GASNDNHYFGYSTPMGYPDPNRFCHFSRDMQRLINNNMGFPKRLNPLFIQYK	119
QY	323	EYTTNDGVTTIANNLSTVQVFSDESYQLPYVLGSAHQCLPPFPADVEMIDQYGLTLN	382
DB	120	EYTTNDGVTTIANNLSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVEMIDQYGLTLN	179
QY	383	NSQAAGRSSFYCLEYFSPQMLRTGNNFTFSYTFEEVPHSSYASQSIDRLMPLIDY	442
DB	180	NSQAAGRSSFYCLEYFSPQMLRTGNNFTFSYTFEEVPHSSYASQSIDRLMPLIDY	239
QY	443	LYLTARTQNSGSAQKDLFSRGSFAGMSVCPKXNLPPCPYRQGVSTKTDNNNSNFP	502
DB	240	LYLTARTQNSGSAQKDLFSRGSFAGMSVCPKXNLPPCPYRQGVSTKTDNNNSNFP	299
QY	503	MFGASKYNNGRESIIPGTAMASHKDDKFFPMGSMWIPGKESAGASNTALDNVMTD	562
DB	300	MFGATKYNNGRDSLVP--AMASHKDDBEKFFPQSGVLIIFGQSEKTNVNIEXKMTD	357
QY	563	EEBIKATNPVATERFQTVAVNFQSSSTDPAQDVHAMGALPGKMWODRIVYLOGPIWAKI	622
DB	358	EEBIGTTNPVATERFQSVSTNLQGRNQAATADVNTQGVLPKGMVWODRIVYLOGPIWAKI	417
QY	623	PHTDGHFHSPLMGFGFLKNPPQQLIKTTPVDPANPAFSAATKFSFTQVSTG	677

Db 418 PHTDGHFHSPLMGFGGLKHPPEQLIKNTVPANPSTTFSAKRASFITQYSTG 412

RESULT 2

552210
coat protein VP1 - muscovy duck parvovirus
N:Alternate names: VP1 protein
C:Species: muscovy duck parvovirus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S52210
R:Zadori, Z.; Erdel, J.; Nagy, J.; Kisari, J.
submitted to the EMBL Data Library, September 1993
A:Reference number: S52209
A:Accession: S52210
A:Molecule type: DNA
A:Residues: 1-732 <ZAD>
A:Cross-references: UNIPROT:O83289; UNIPARC:UPI00000F6C5D; EMBL:X75093; NID:g609091; PID
A:Experimental source: strain FM
C:Genetics:
A:Gene: VP1
A:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 54.4%; Score 2168.5; DB 2; Length 732;
Best Local Similarity 55.2%; Pred. No. 2.6e-133;
Matches 418; Conservative 90; Mismatches 186; Indels 63; Gaps 13;

Qy 9 DWLENDLSEGLREWMDLKQAPKPYANOQK-----DGRGLVLPQYKYLGFN 57
Db 10 DWYE-----TAAASMRHLKAGAPKPKSNQOSVSTDRKPKQKNNRGVFLPKYKYPGN 65
Qy 58 GLDKSEPVNAADALEHDKAYDQOLKAGDNPLYLNHADAFOERLOEDTSFGGNGRA 117
Db 66 GLDKSEPVNAADALEHDKAYDQOLKAGDNPLYLNHADAFOERLOEDTSFGGNGRA 125
Qy 118 VFQAKRVLLEPLGLVEBGAKTAPGKKRVEQSPQSDSSGIGKTGOOPAKKRLNFGQTA 177
Db 126 VFQAKRVLLEPLGLVEBGAKTAPGKKRVEQSPQSDSSGIGKTGOOPAKKRLNFGQTA 170
Qy 178 DSEVDPDPOPLG-----EPATPAVAGPTTMASSGGAPADNNEGADGVGNASG 226
Db 171 --ENSPSPNSGSGEASAAATBGSBEVPAAP-----NMAEGSGAGADBSAGADGVGNASG 222
Qy 227 NMHCSTVLGDRVITTSRTMALPTYNHLYKQISSASGASNDNHGYGSPKMGYPEN 286
Db 223 NMHCSTVLGDRVITTSRTMALPTYNHLYKQISSASGASNDNHGYGSPKMGYPEN 281
Qy 287 RFHCHFSRDMQRLINNMGFRPKRLNFKLENIQVKEVTTNDGVTIANNLTSTQVPSD 346
Db 282 RFHCHFSRDMQRLINNMGFRPKRLNFKLENIQVKEVTTNDGVTIANNLTSTQVPSD 341
Qy 347 SEYQLPYVLGSAHQCLPPFPADVMIPOYGYLTN--NSQAVGRSSPYCLEYFPSPQM 403
Db 342 NEHQPYVLGSAHQCLPPFPADVMIPOYGYLTN--NSQAVGRSSPYCLEYFPSPQM 401
Qy 444 LRTGNFTFSYFEFVPPHSSYAHOSLDRLNPLIDQVLYLTNTQSSGAQKDLF 463
Db 442 LRTGNFTFSYFEFVPPHSSYAHOSLDRLNPLIDQVLYLTNTQSSGAQKDLF 455
Qy 464 SRGSPAGMSVQPKNMLPGPCYRQORVSKTK--TDNNNSNFTWTGASKYVLNGBISLNG 521
Db 466 KKAUVGAPGAMGRNMLPGPKLDQVRAVSGGTDYAMNSIKSNKVPFLKQREYLDQPG 515
Qy 522 TAMASHKODEDKFPFMSGVMIFGKE--SAGASNTALDNVMTDEBEIKATNPVATERPQT 579
Db 516 PAATHTEDQASSVPAQNIIGIAKDPYRSGSTLAGISIDIMTDEOEIAPTNVGMRPYGL 575
Qy 580 VAVNQSSTDPATGVDHAMGALPGMWQDRDVLVQGPIMAKIPHTDGHFHSPLMGCG 639
Db 576 TVTNQONTTAAATNMLEVLGALPGMWQDRDVLVQGPIMAKIPHTDGHFHSPLMGCG 635
Qy 640 LKNPPQOLIKNTVPANPSTTFSAKRASFITQYSTGQCTVEWVWELRKENSKRMPVEI 699

Db 636 LHNPPQOLIKNTVPANPSTTFSAKRASFITQYSTGQCTVEWVWELRKENSKRMPVEI 695

Qy 700 QYTSNYAKSANVDFTVNNGLYTEPRPIGRTYLLRPL 736
Db 696 QYTSNYAKSANVDFTVNNGLYTEPRPIGRTYLLRPL 732

RESULT 3

WCPV25
coat protein VP1 - bovine parvovirus
N:Contains: coat protein VP2
C:Species: bovine parvovirus
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A26104
R:Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.
J. Virol. 60, 1085-1097, 1986
A>Title: Complete nucleotide sequence and genome organization of bovine parvovirus.
A:Reference number: A26104; MUID:87061184; PMID:3783814
A:Accession: A26104
A:Molecule type: DNA
A:Residues: 1-673 <CHE>
A:Cross-references: UNIPROT:P07297; UNIPARC:UPI0000127D77; EMBL:M14363; NID:g333454; PID
A:Superfamily: parvovirus coat protein
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:138-673/Product: coat protein VP2 #status predicted <VP2>

Query Match 20.0%; Score 798.5; DB 1; Length 673;
Best Local Similarity 28.6%; Pred. No. 3.7e-44;
Matches 214; Conservative 105; Mismatches 275; Indels 155; Gaps 21;

Qy 30 PKPKANQKODDGLVLPQYKYLGFNGLDKSEPVNAADALEHDKAYDQOLKAGDN 89
Db 3 PTNKANSK-----GLTPGYNYLGPFRNSLFAAGPYNKADAAKRDGYGDDLKQGNP 57
Qy 90 YLRYNHADADEFQERLOEDTSFGGNGRAVFGAKRVLLEPLGLVEBGAKTAPGKKRVEQSP 149
Db 58 YLRYNHADADEFQERLOEDTSFGGNGRAVFGAKRVLLEPLGLVEBGAKTAPGKKRVEQSP 111
Qy 150 PQEDSSGIGKTGOOPAKKRLNFGQCTDSSVDPDPOPLGSP--PATPAAGPTTMASSG 207
Db 112 LYPARSNKGAKKARREAPSTSNQONNEVSNIDIPDEAGNPILATRSVVGSGSGVGGG 171
Qy 208 GAPADNNEGADGVGNASGGMHCSTVLGDRVITTSRTMALPTYNHLYKQISSASGTA 267
Db 172 -----RGSGSYGYSIGGTTGTTISENIVTKRQFICDKNGHLYS--EVLNIGD 222
Qy 268 SNDNHGYGSPKMGYPENRFHCHFSRDMQRLINNMGFRPKRLNFKLENIQVKEVTTN 327
Db 223 TAHROY-ALTTPMGEYFNQSSHSFSPMDQHLVNDYERFRPKAMIVAVNMLQIKQIMTD 281
Qy 328 DGVTTIANN-LTSTVOVPSDEYQLPYVLGSAHQCLPPFPADVMIPOYGYLTN----- 381
Db 282 GAMGVVYNNDLTAGMHFLCDDDHRYVYQNHPMDOCEMLNSIWEIYQVAYIIPAPISVV 341
Qy 382 --NNGSOAVGRS-----SFYCLEYFPSPQMLRTGNFTFSYFEFVPPHSSYAHOSLDRLN 435
Db 342 DNNTTNYEELHLKGVPLYLHNSDHYLRNG-----RIV 376
Qy 436 NPLIDQVLYLTNTQSSGAQKDLFSGSP-----AGMSVQPKNMLP 480
Db 377 RIVQLWRLRMDRQHNHIAHSDVQSTGQKQKILLIQRKQPKRQFQNALARTSNMS 436
Qy 481 GPCTRQQRVSKTKIDNNNSNFTWTGASKYVLNGBISLNGEESIINPTMAASHDDEDDKFPMSGV 540
Db 437 GP-----GIAGTGNATLQTSAGALVMTVNTGAD---VSGV 470
Qy 541 M-----IFGKE-----SAGASNTALDNVMTDEBEIKATNPVATERPQT 575
Db 471 RAVRVGSTDPIYGGQCPESDLRLRYASAAEQQNILEN-----AAAH 516
Qy 576 RFGTVAVNQSSTDPATGVDHAMGALPGMWQDRDVLVQGPIMAKIPHTDGHFHSPLMG 635
Db 517 TTTREARLKLITGSGAGADGVKEMMMLPQNMWDASPISRVNDIWKVPRVNRKTLTDTPD 576

Qy 636 GGFGLKPPPOILLK--NTVPANPPAEFSATKFPSTTOYSTGCVSVIEIEMELOKENSK 693
Db 577 GSIPMSHPPTGFIKILARIPVGNCD-----SFLNIYVTGQVSGEVWEVEKRGTK 627
Qy 694 RKNPEVQTSNKAASND-FIVDNNGLY 721
Db 628 NMRPEYMS--ATNMSVDATYINNAGVY 653

RESULT 4

vcpv19
coat protein VP1 - parvovirus B19 (strain Au)
C:Species: parvovirus B19
A:Note: host Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A24299
R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A:Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr
A:Reference number: A24299; MUID:86200451; PMID:3701931
A:Accession: A24299
A:Molecule type: DNA
A:Residues: 1-781 <SHA>
A:Cross-references: UNIPROT:P07299; UNIPARC:UPI0000127D7C; EMBL:M13178; NID:G333375; PID
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 15.1%; Score 600.5; DB 1; Length 781;
Best Local Similarity 26.6%; Pred. No. 3.5e-31; Indels 107; Gaps 20;
Matches 182; Conservative 95; Mismatches 300;

Qy 47 LFGYKYLGFPGNLGDKPEVNAADAALAEHDKAYDOOLKAGDNPYLRYNHADAEPERLOE 106
Db 125 LFGTYVVGNGELQNGPQSAVDSARINDPYSQAKLGINPYTHMTVADELLKNIKN 184
Qy 107 DTSPGNGIGRAVFOAKKRVLEPLGLVEBQAKTAPGKKRVEQSPQEPDSSSGIGTKGQOP 166
Db 185 ETGFQAOVAKDYF-----TLKGAAPVAHFQ-----GSLPEVP 217
Qy 167 AKKRLNFGQGTGSESVDPQPLGEPPATPAVQPTTMASGGAPMADNNEGADGVGNAG 226
Db 218 ATNA-----SEKYPMSITSVASASATGA-----GGG-----GSNSVKS 250
Qy 227 NNHCDSTWLGDRTITSTRTMALPTYNHLKYQISASTGASND-----NHIFY 276
Db 251 MMSGATFESANSVTCTFSQFLIPYRPEHHKYFSPASCSNASKAKVCTISPIKMY 310
Qy 277 STPMWGFDENRPHCHSPRDWQRLINNMGFRPKRLNFKLFNIQVEVT--TNDVTTTA 334
Db 311 STPMRYLDENALNLFSPLEFQHLIENYGSIAFDALTITISEIAVDVLDKGGGV--QVT 369
Qy 335 NNLTSTVQFSDSEYOLPYVLGSAHGCCLPPPADVFMIPQCYLLNN--GSAVG----- 389
Db 370 DSTTGLCMVLVHEHYRYPVLLGGQDTLAPELRIWYPPQYAYLLVGVGNTQISGDSK 429
Qy 390 -----RSSFYCLEYFSPQMLRTGNNTFSYTFEEVPHSSYAHQSGLDRLMPLIDQY 444
Db 430 KLAESASATYVLEHSPQLLGTGCTASMGYKPPRYPEMLGSCQFYEMYNL---YGS 486
Qy 445 YLNRTNOSGSAONKDLFSRSGPAGMSVQPKWMLFGPCYRQORVSKTKTDNNNSFTMT 504
Db 487 RLGVPRDLGGDPKFRSL-----THEDAIOPQNFMGPLVNSVSTKEGDSSTNGAKAL 541
Qy 505 GASKYLNREGSLING--TAMASHKDEDEKFFPMGSMVIFGKESAGASTALDNV----- 558
Db 542 GLSTGTSQNTRLSLRGVPSQPYHMDTDKVTGINAISHGQTTYG---NAEDKEYQGY 598
Qy 559 --MITDEEIKATNPVATERFGTAVANFOSSSTDPAQDVHAGALPGMWQDRDYVLQ 616
Db 599 GRPNEKEQIKOGLNMHTY-----FRPKGQOYTDQLE--RLWVGVSVMNRALHYES 651
Qy 617 PIWAKIPHTDGHFHS--PLMGSGFLKNPPQILLKQTPVNPANPPAEFSATKFPSTTOYS 675

Db 652 QLMSEKIPRLDSFRTQFALGKGLHQPPPIFLK--ILPQSGPIGIGIKSGITTLVGYA 709
Qy 676 TGQVSVIEIEMEL--QKENSKRNP 698
Db 710 VGIHTVWTWTFKLGPRKATGRRNPQ 733

RESULT 5

A60006
coat protein VP1 - porcine parvovirus (strain 90HS)
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: A60006
R:Sakurai, M.; Nishimori, T.; Ushimi, C.; Nakajima, H.
Virus Res. 13, 79-86, 1989
A:Title: Nucleotide sequence of capsid protein gene of porcine parvovirus.
A:Reference number: A60006; MUID:89319168; PMID:2750278
A:Accession: A60006
A:Molecule type: DNA
A:Residues: 1-729 <SAK>
A:Cross-references: UNIPROT:P33484; UNIPARC:UPI0000127C6E
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein; glycoprotein
F:151-723/Product: coat protein VP2 #status predicted <VP2>
F:172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #statd

Query Match 12.1%; Score 482; DB 1; Length 729;
Best Local Similarity 25.7%; Pred. No. 1.7e-23;
Matches 197; Conservative 112; Mismatches 314; Indels 144; Gaps 35;

Qy 43 RGLVLPGYKYLGFPGNLGDKPEVNAADAALAEHDKAYDOOLKAGDNPYLRYNHADAEPER 102
Db 9 RGLTLPGYKYLGFPGNSLDQGEPTNPSDAAAEHDEAYDKYIKSGKNPYFYFAADEKFLK 68
Qy 103 RLQEDTSGCNLGRAVFOAKKRVLEPLGLVEBQAKTAPGKKRVEQSPQEPDSSSGIGK- 161
Db 69 ETEHAKDYGKIGHYFFRA--KFAFRP--KLSTDSPTTSQPEVRSRPAKHGSRPGRK 125
Qy 162 -----TSQOPAKKRLNFGQGTGSESVDPQPLGEPPATPAVQPTTMASGGAPMA 212
Db 126 PAPRIHFINLAKKRAKKTNSNTSMSENVEHNPINNGETL--SATNGSGGGGGG---- 181
Qy 213 DNNEGADGVNAGSNW--COSTWLGD--RYTTSTRTMALPTYNHLKYQIS--SAST 265
Db 182 -GGRGAGGVGVSATGSPNNQTEFOYLEGIVRTAHASRLIHLMPEHEHYKRIHYLNSFS 240
Qy 266 GAS-----NDNHFGYSTWGYDFRPHCHSPRDWQRLINNMGFRPKRLNFKLFNIQV 321
Db 241 GVAQGMVDDAHTQWVTPMSLIDANAGVWENPADWQLISNMTEINLVSPQBIENVYL 300
Qy 322 KEVT--TNDVTTIANNTSTVQFSDSEYOLPYVLGSAHGCCLPPPADVFMIPQCY 378
Db 301 KTTTSATSPRKIKIYNNDLTASLMALDTNNLTPTTPAARSETIGFYWMLTKPQOYX 360
Qy 379 L-----TLNNGSAQVGRS-----SFYCLE--YFSPQMLRTGNMF--TSTSYTE 417
Db 361 YLSTCRNLNPPYTGQSQDITDSIQTGLSDIMFYTIENAVDIIHLRTDDESTGIYHD 420
Qy 418 EYFPHSSYAHQSGLR-----LMNPLI--DOYLYLNRTNONGSGAONKDLFSRSG 467
Db 421 TRPL--KLTHSQTRSLGLPKLLTEPTBGDQHGTLTPAANTRGHNQTMNSTYTEAT 478
Qy 468 PAGMSVQPKWMLPGPCYRQORVSKTKTDNNNSFTMTGASKYLNREGSLINP--GTAMA 525
Db 479 ----AIRP-----AQGVNPPYNNFEVSNQGP-----LTPVPTADT 512
Qy 526 SHKDEDEKFFPMGSMVIFGKESAGASTALDNVMTDEEIKATNP-----VATERP 579
Db 513 QVNDDE-----PAGAIRFTWQYHQLTTS-----SOELERYTFPOSKCGAPRKOQFNO 562
Qy 580 VA--VAFQSS-----TDPATG--DVHAGALP-----GMWQDRDQVYLQGPWAKIPHT 625

```
Db 563 QAPLNLENTNNGTLLPSPDPIGKPMHFMTLNTYGPALTALNNTAPVPNGQIWDKELDT 622
Oy 626 D--GHFHSPLMGCGFLK-NPPQILIKNTVPANPAPFSA-TRFASFITYQSTGVSV 681
Db 623 DLKPRLLH---VTAPFVCKNNPPGQLFVKIAP---NLTDFFNADSPQOPRIITYSNFMWKG 676
Oy 682 EIEWELQENSKRMNPEVOYTSNYAKSANVDFTVDNNGLYTEPRPIG 728
Db 677 TLTFPAKRSSNMNPIQOHTT-----TAENIGNYI-PTNIG 712

RESULT 6
VCPVNA
coat protein VP1 - porcine parvovirus (strain NADL-2)
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: B33743; D48472
R:Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.
V:Virology 173, 368-377, 1989
A>Title: Nucleotide sequence analysis of the capsid genes and the right-hand terminal pe
A:Reference number: A33743; MUID:90085785; PMID:2596019
A:Accession: B33743
A:Molecule type: DNA
A:Residues: 1-729 <VAS>
A:Cross-references: UNIPROT:P18546; UNIPARC:UPI0000127C70; GB:M32787; NID:g332983; PIDN:
R:Bejteron, J.; Meneses, J.; Tjissen, P.
V:Virology 197, 86-98, 1993
A>Title: Genomic organization and mapping of transcription and translation products of t
A:Reference number: A48472; MUID:94025614; PMID:8212598
A:Accession: D48472
A:Molecule type: DNA
A:Residues: 11-729 <BER>
A:Cross-references: UNIPARC:UPI0000174965
A:Experimental source: NADL-2, ATCC VR-742
A:Notes: sequence extracted from NCBI backbone (NCBIN:138789, NCBI:P138794)
A:Genetics:
A:Introns: 10/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein, glycoprotein
F:151-729/Product: coat protein VP2 #status predicted <VP2>
F:32,172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 12.1%; Score 482; DB 1; Length 729;
Best local similarity 25.5%; Pred. No. 1.7e-23;
Matches 197; Conservative 111; Mismatches 310; Indels 154; Gaps 35;

Oy 43 RGLVPGYKYLGPFGNGLDKGEFVNADAALAEHDKAYDQOLKAGDNPYLRYNHADAFOE 102
Db 9 RGLTLPGYKYLGPFGNSLDGSEPTNSDAAAKHDAVYKIKSGNPFYFSADEKTIK 68
Oy 103 RLQEDTSGGNLGAIVPAQKKRVLEPLGLVEGAKTAPGKKRPVQSPQEBDSSSGICK- 161
Db 69 ETEHAKDYGKIGHYFFRAKRAFAPKL---SETDSPTTSQOPEVARSBRKPHGSKRPGR 125
Oy 162 -----TGOQPAKKRLNFCGTGDSVPDPOGLGEPRAIPAIVGPTTASGGGAMA 212
Db 126 PAPRIHFINLAKKAKGTSNTNSMSSEVQHNINAGTEL-SATGNESGGGGG--- 181
Oy 213 DNNEADGVGNASG--NMHCSTWLGD--RVITSTRTALPTYNHLYKOIS--SAST 265
Db 182 -CGRAAGGVSTGFENNQTEFOYLOBLVITTAASRLIHNMHEHETVKRIHLYNES 240
Oy 266 GAS---NDNHFGYSTWGYDFNRFCHFSRPRDWRLLNNWGFRRRLNFKLFNIQV 321
Db 241 GVAQGVQDDAHTQMTFMSLIDANAWGVFNPADQULSNMTEINLVSPQEIFNVTL 300
Oy 322 KEVT---TNDGYTTIANNTLTVOYFSDSEYOLPYVLSAHQGCPRPADVFMIPQYX 378
Db 301 KITTSATSPPTKIKYNNDLTASLMVALDTNNTLPTTAPARSETTGFPWMLPTKPTQYX 360
Oy 379 L-----TLNNGSAVGRS-----SPCYCLE-YFPGQLRIGGNP-TESYTFE 417
```

```
Db 361 YLSICRLNAPPTYGQEQQTDSIQTGASHDIMFYETIENAVPIHLLRTGDEFSIGIYHFD 420
Oy 418 EVPHSSYASQSLDR-----LNMPL--DOYLYLNRTONQSGAQNKLFFSRGS 467
Db 421 TKPL--KLTHSWQTNRLGLRPKLLEPTEGDOHPECTLPAAATRKRYHOTINNSTYBEAT 478
Oy 468 PAGSVQPKMWLDPQCYROQVRSKTKTDNNNSNETWGAQKYNLNGRESIINP--GTAMA 525
Db 479 ----AIRP-----AQVGYNTPYNMEVINGRPF-----LTPVLPADT 512
Oy 526 SHKQEDKFPFGSGWMIFGKESAGASNTALD--NVMITDEBEIK--ATNP-----VAT 574
Db 513 QYNDEBN-----GAIRFTMDYGHGLTTSQELERYTFPQSKGRAPK 557
Oy 575 ERFGTVA-VNFGSS-----TDPATG--DYHAMGALP-----GMVQDDVYLCQPIWA 620
Db 558 QQFNOQAPLNEINTNNGTLLPSPDIGKSNMHEFNTLNTYGPALTALNNTAPVPNGQIWD 617
Oy 621 KIPTD--GHFHSPLMGCGFLK-NPPQILIKNTVPANPAPFSA-TRFASFITYQST 676
Db 618 KELDIDLKPRLLH---VTAPFVCKNNPPGQLFVKIAP---NLTDFFNADSPQOPRIITYSN 671
Oy 677 QGVSVIEIEWELQENSKRMNPEVOYTSNYAKSANVDFTVDNNGLYTEPRPIG 728
Db 672 FWMKGLTLTFPAKRSSNMNPIQOHTT-----TAENIGNYI-PTNIG 712

RESULT 7
VCPVIF
coat protein VP1 - feline panleukopenia virus
N:Contains: coat protein VP2
C:Species: feline panleukopenia virus, FPLV
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A03701
R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
V:Virology 155, 574-587, 1985
A>Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parv.
A:Reference number: A03697; MUID:85265017; PMID:2991581
A:Accession: A03701
A:Molecule type: DNA
A:Residues: 1-727 <CAR>
A:Cross-references: UNIPROT:P04864; UNIPARC:UPI0000127D55; EMBL:M10824; NID:g333474; PID:
A:Genetics:
A:Introns: 11/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.1%; Score 444.5; DB 1; Length 727;
Best local similarity 23.7%; Pred. No. 4.6e-21;
Matches 187; Conservative 117; Mismatches 225; Indels 189; Gaps 35;

Oy 43 RGLVPGYKYLGPFGNGLDKGEFVNADAALAEHDKAYDQOLKAGDNPYLRYNHADAFOE 102
Db 10 RGLVPGYKYLGPFGNSLDGSEPTNSDAAAKHDAVYKIKSGNPFYFSADEKTIK 69
Oy 103 RLQEDTSGGNLGAIVPAQKKRVLEPLGLVEGAKTAPGKKRPVQSPQEBDSSSGICKT 162
Db 70 QTKDATTMGKIGHYFFRAKRAIAVLTDPDHSRTPRK-PTKRSRPPHIFINLAK- 127
Oy 163 GQOPAKKRLNFGQGDSESVP-----DPOPLGEPPA--TPAAVGPPTTASGGGAPADNNE 216
Db 128 -----KKKAGAGQYKRDNQAPMSGAVQPDGQGAIVNERNRTGSGNSGGG-----CG 176
Oy 217 GADGVGNASGNM-----CDSTWLGDRTVITSTRTALPTYNHLYKOI-----SS 262
Db 177 GSGGVGISTGFENNQTEFKFLENGAV--ETRANSRLVHLNMPSESEYKRVVANNMDKTA 234
Oy 263 ASTGASDNHFGYSTWGYDFNRFCHFSRPRDWRLLNNWGFRRRLNFKLFNIQYX 322
Db 235 VKGNMALDDTHVQIVTPMSLVADANAWGVFNPADQULSNMTEINLVSPQEIFNVVLK 294
```



```

Db      70 QTKADKMGKIGHYFFRAKKAIAFLVLTDPDHPSTSRPTK-PTKRSKPPHIFINLAK- 127
Qy      163 GQPAKXKLNTGQTDSESV-----DQPLGEPRA--TPAVGPTTMAAGGAPMADNE 216
Db      128 -----KKKAGAGQVRNDLAPMSDGAVQPDGQPAVRNERATGSGNGSGGG-----GG 176
Qy      217 GADVGNSAGMWH-----CDSTWLGDRVITTTSTRTVALPTYNHLYKQI-----SS 262
Db      177 GSGGIGISTGTFNNQTEFKFLENGMV--EITANSRLVHLNMPSEENKRVVYNNMDXTA 234
Qy      263 ASTGASNDNHFGVISTPMGYDFENRFFCHFSRDMQRLINNNMGPRKPLNFKLFNIOVK 322
Db      235 VKGNMALDIDHVOIYTPWMSLYDANAMGWFFNPGMQLIVNTMSSELHVSFPEGLFNVVLK 294
Qy      323 EYT-----TNDGVTTIANNTLSTVQVPSDEYOLPYVLGSAHQCL-----PPFPA----- 368
Db      295 TVSSATQPPKRVYNNDLASLMAVLDSSNNTMFPAPAMRSETLGFYPMKPTIPIPMKYY 354
Qy      369 ---DVFMIPQYGYLTLNNGSQAVGRS-----SFYCLE-YPPSQMLRTGNMF-TFS 413
Db      355 FQWDRTLIPSH-----TGTSCTPTNVVHGTDPDQVGYTTIENSVPVHLRTGDEFATGT 408
Qy      414 YTFEEVPHSHSYAHSQSILDRLMN--PLIDQYLYINRTONOGSAQNKDLFSGSPAGM 471
Db      409 FFFDCKP--CRLLHTMOTNRALGLRP-----PLNSLPQSRGATNFGDIGVQDCKRGV 459
Qy      472 S-----VQPK-MLPGCYRQQRYSK-----TKTDNNNSNFTMT 504
Db      460 TOMGNTDYITTEAKTMRPAVGVSAHYSEASTQCPPTPIAAGGAGQTDENQA---D 516
Qy      505 GASKY---NLNGRESIIINGT---AMASHKDEDEKFFPMGCVMLFGKESAGASNTALD- 556
Db      517 GQPRVAFGRHQKQKTTTGTGETPERFYIAHQDT-----GAYPEGDWITQINNF 563
Qy      557 NMVITDEEIKATNPVATERTFGTVAVNFGSSSTDATGDAVHAGMLPGMWODRQVYLOG 616
Db      564 NLPVNDVNLPTDIDIG---GKTGINY--TNIPTYGPTLAINVP-----FYVING 610
Qy      617 PIWAKIPHTDGHFHFSLMGFGLK-----NPPQILIKNTVPAN---PPA 660
Db      611 QIWDKEPDTD-----LKRLLHVNAPFYCQNNCPQQLVTKVAPNLITNYYDDPA 657
Qy      661 EFSATKFAFITQYSTGQVSEIEMLEKENSKRNNPEVQYTSNVAKSANVDFTVNNGL 720
Db      658 SANMR-----LYTSDFMWKGKLVFKAKLRASHTNPIQOMSI-----VDNQFN 703
Qy      721 YTEPRPIG 728
Db      704 YV-PNNIG 710

RESULT 10
vcvpp
coat protein VP1 - porcine parvovirus
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: B33302
R:Ranz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casal, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A:Title: Porcine parvovirus: DNA sequence and genome organization.
A:Reference number: A33302; MUID:90010964; PMID:2794971
A:Accession: B33302
A:Molecule type: DNA
A:Residues: 1-723 <RAN>
A:Cross-references: UNIPROT:P18546; UNIPARC:UPI0000174964; EMBL:DD0623
C:Genetics:
A:Introns: 10/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:145-723/Product: coat protein VP2 #status predicted <VP2>
Query Match 10.7%; Score 426; DB 1; Length 723;

```

```

Best Local Similarity 24.9%; Pred. No. 7.3e-20;
Matches 190; Conservative 109; Mismatches 308; Indels 156; Gaps 36;

Qy      53 LGPENGIDKEPVAADAALAHDXADQOLKADNLYLRNHADEFOELQSDTSFG 112
Db      13 LGPNSLIDQEPNPDAALAKHDEADYDKYSKNTFYPSADEKFKETBAKDIYG 72
Qy      113 NLGRAVQAKKRVLEPGLVEGAKTAPKRRPYEQSPQESDSSGIQK-----T 162
Db      73 KIGHYFFRAKRAFAPKL---SETDSPTTQPEVRARBRPKHPSGKPPQKRAPRAPHIFNL 129
Qy      163 GQPAKXKLNTGQTDSESVPDQPLGEPRA--TPAVGPTTMAAGGAPMADNEGADV 221
Db      130 AKKAKAGTSMNSMSSENVQHNPI--AATLSATGNSGGGGGGG---GGRGAGV 183
Qy      222 GNAAGMNH---CDSTWLD---RVITTSRTVALPTYNHLYKQI-----SSASTGAS-ND 270
Db      184 GVSSTFSNNQTEFOYLGGLVRYTAHASRLIHLMPHEHYTKRIHVLNSBSGSAQOMYVD 243
Qy      271 NHYFGYSTPMGYDFENRFFCHFSRDMQRLINNNMGPRKPLNFKLFNIOKVEYV---TN 327
Db      244 DAHTQMTPLMSLIDANAMGWFFNPDQMLISNNMTINLVSFEQALFNVLKTTESATS 303
Qy      328 DGVTIANNTLSTVQVPSDEYOLPYVLGSAHQCLPPFPADVMIPOYGYL----- 379
Db      304 PPTKIYNNDLASLMAVLDNTNLTPTPAARSETLGFYPMKPTQYRYLSCIRNLN 363
Qy      380 ---TLNNGSQAVGRS-----FYCLE-YFPQMLRTGNMF-TFSYTFEEVPHSSYA 426
Db      364 PPTYGQSQPNRNRLNTRLSHDIMFYTTENAVPHLRLTGBESTGTYHFDTKPL--KLT 421
Qy      427 HSQSLDR-----LNNPLI--DOYLYLINRTONOGSAQNKDLFSGSPAGMSVQPK 476
Db      422 HSWQNTSLGLRPKLLEPFTTEGQHPGTLPAANTBKGYNQITNSYBEAT---AIRP- 476
Qy      477 NMLPGCYRQQRVSKTKTDNNNSNFTWTGASKYVNLNGRESIINP--GTAMASHKDEDEK 534
Db      477 -----AQVGNTPYNMFYSNGPF-----LTPIVPTADQVYDDEPN- 514
Qy      535 FPMGCVMLFGKESAGASNTALD---NMVITDEEIK--ATNP-----VATERGTVAV- 582
Db      515 -----GALRFTMDYOHGHLTSSQELERTFPPQSKGRAPRQQRNOQAPL 560
Qy      583 NFGSS-----TPRATG--DVHAMGALP-----GMWQDRDYVLOGPIWAKIPHTD--G 627
Db      561 NLENTNNTLLPSPIGSKSMHFMNTLNTYGPLTALNNTAPVPPNQGIWKEIDTDCKP 620
Qy      628 HFHPSPLMGFGFLK-NPPQILIKNTVPANPAPFSA--TKFASFITQYSTGQVSEIEM 685
Db      621 RLH---VTAPVVCQKNPFGQLFVKIAP---NLTDFFNADSPQOPRIITYSNFMWKGTLTF 674
Qy      686 ELQKENSKRNNPEVQYTSNVAKSANVDFTVNNGLYTEPRPIG 728
Db      675 TAKRRSSNMNPIQOHTT-----TAENIGKYL-PTNIG 706

RESULT 11
vcvpm
coat protein VP1 - minute virus of mice (strain MMV1)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: B23008; B29510
R:Sahl, R.; McMaister, G.K.; Hart, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Accession: B23008
A:Molecule type: DNA
A:Residues: 1-718 <SAH>
A:Cross-references: UNIPROT:P07302; UNIPARC:UPI0000174963; EMBL:X02481; NID:g60918; PIDN
R:Abell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A:Title: DNA sequence of the lymphotropic variant of minute virus of mice, MMV(1), and c

```


RESULT 13

VCPVCD

coat protein VP1 - canine parvovirus (strain CPV-d)

N:Contains: coat protein VP2

C:Species: canine parvovirus, CPV

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: A31163

R:Partiuh, C.R.; Aguadro, C.F.; Carmichael, L.E.

Virology 166, 293-307, 1998

A:Title: Canine host range and a specific epitope map along with variant sequences in th

A:Reference number: A31163; MUID:89020796; PMID:3176341

A:Accession: A31163

A:Molecule type: DNA

A:Residues: 1-737 <PAR>

A:Cross-references: UNIPROT:P17455; UNIPARC:UPI0000127D7A; EMBL:M23255; NID:g333467; PID

C:GeneInfo:

A:Inserts: 26/3

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

F:584-737/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.2%; Score 407.5; DB 1; Length 737;

Best Local Similarity 22.8%; Pred. No. 1.2e-18;

Matches 181; Conservative 112; Mismatches 301; Indels 201; Gaps 32;

```
QY 42 GRGLVPGYKYLGPFGNGLDKGRVNAADAALEHDKADNPKYLRNHADAFQ 101
DB 19 GGGKDLITYKYLGPFGNSLDQEPPTNPSDAKHEDEATYALRSKNPILYFSPADQRFI 78
QY 102 ERLQEDTSPFGNIGRAVFOAKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSGIGK 161
DB 79 DQTKAKOMGKIGHYFFRAKKAIALPVLIDTPDHSTSPRK-PTKRSKPRHIFITLAK 137
QY 162 TQOQAKKRLNLCQGTDSRSP---DPOPLSEPPA--TPAAVGTMTASGGGAMADNN 215
DB 138 -----KKKAGAGQYKRDNLAPMSDGAAYQPDGQGPVAVNERATGSGNGSGGG-----G 185
QY 216 EGADGVGNASGMNH-----CDSFTMLGDRVITTSRTMALPTYNHLYKQI-----S 261
DB 186 GSGGAGGISTGFNNOTERKFLENGKV--ETIANSRLVHLMPRESENRYRVVANNMOKT 243
QY 262 SASGTASNDNHFGYSTPWGYPDFNRFCHFSPPDMQRLINNMGFRPKRLNFKLFNIQV 321
DB 244 AANGMAALDDHIAQVLTPLMSLVANAMGWFMFGMQLVNTMSLHLVSFQELFNVL 303
QY 322 KEVT---TNDGVTTANNLTSTVOVFSSEYQLPYVLGSAHOGCL-----PPPPA----- 368
DB 304 KTVSESATOPPTKYVNNNDLTASLMAVLDSSNNTMPTTPAAMRSETLGFPYPMKPTIIPTRRY 363
QY 369 ----DVFMIPQYGYLTLNNGSAVGRS-----SPCYCLE-YFPSCMLRTGNNF--TF 412
DB 364 YQWMPRTLLPSH-----TGTSGTPTNIYHGTDPDQVQYTTIENSVPVHLTGTGBFATG 417
QY 413 SYTFEEVPRHSSVYASQSLDRLMN--PLIDQYLYLNTONOGSQAONKDLFSRGSBAG 470
DB 418 TFFPFOCKP--CRLLTHWQTNRAIGLPR-----FLNSLPQSBGATNREDI-----G 460
QY 471 MSVQPRKMLPGPCYRQQRVSKTKTNNNSNF-----TWTKASKTNLNGRESIINP- 520
DB 461 V-----QODKRGVTOGNTNYITTEATIMRAVEGVSAPIYSSFEASTOGPF 506
QY 521 ----GTAMASHKDDKXFFPMGSVMI FGKESAGASNTALD----- 556
DB 507 KTRPIAAGGAGQDDEQAADGNPRYAFGHOQOKTTTGTGTERDRTFYIAHODTGKYPED 566
QY 557 -----NMTTDEERIKATNPATERFGTVAVNFQSSSTDPATGVDHAMGALPGWVMD 609
DB 567 WTONINFNLPTVNDNLPLTDPDG--GKTGINY--TWIENYGLTALNNVP----- 614
QY 610 RUVYLQGPYIMAKIPTHGTHFHSBPLMGFGK-----NPPQILIKTTPVPA 656
DB 615 -FVYPRGQITMDEKFDTD-----LKRPLHVNAPFVCONNCPQLFVKAENLT 660
```

RESULT 14

VCPV2

coat protein VP1 - parvovirus H1

C:Species: parvovirus H1

A:Note: host Homo sapiens (man)

C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004

C:Accession: A03659

R:Rhode II, S.L.; Paradiso, P.R.

J. Virol. 45, 173-184, 1983

A:Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid

A:Reference number: A03655; MUID:83112183; PMID:6823009

A:Accession: A03659

A:Molecule type: DNA

A:Residues: 1-722 <RHO>

A:Cross-references: UNIPROT:P03136; UNIPARC:UPI0000127D7D; EMBL:X01457; EMBL:J02198

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

Query Match 10.1%; Score 404; DB 1; Length 722;

Best Local Similarity 23.3%; Pred. No. 2e-18;

Matches 175; Conservative 108; Mismatches 290; Indels 178; Gaps 30;

```
QY 57 NGDLKGEVNAADAALEHDKAVDQOLKAGDNPYLRYNHADAEQERLQEDTSPFGNIGR 116
DB 12 NSLDQEPPTNPSDAKHEDEATYQIKSGNRPILYFSPADQRFIDQTKAKOMGKYGKH 71
QY 117 AVFOAKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSGIGKTQOQ--PA----- 167
DB 72 YFFRTKAFAPKLSTDS-----PG-----TSGVSRPKRTKRPAHIFVQ 112
QY 168 -----KRLNFCQGTDSRSPDQPLSEPPA--TPAAVGTMTASGGGAMADNNEGADGVN 223
DB 113 ARAKKASALAAQQRITLTMSDGTETNPDTGIANAVERSADGGS--SGGGSGGGGIGV 171
QY 224 ASGMNHCDSTW--LGDRIITSTRMAL-----PYNNHLYKQISSASTGA 267
DB 172 STGYNDQTTKFLGDDGVEITTAASRLHLGMPRENYCVYTHANNQTTGHTKIKYKGM 231
QY 268 SNDNHFGYSTPWGYPDFNRFCHFSPPDMQRLINNMGFRPKRLNFKLFNIQYKEVTT- 326
DB 232 AYDTHQQLW--TPMSLVANAMGWFMFGSDMQFQIONMESLNLDSLQELFVVVVKTVTEQ 290
QY 327 ----NDGVTTANNLTSTVOVFSSEYQLPYVLGSAHOGCL-----PPPAD--VFMI 373
DB 291 QGAGODALKVYNNNDLTACMVALDSNNTLPTTPAAGTSETLGFPYMKPTAPAPRYVFM 350
QY 374 POYGYLTLNNGS-----AVG-----RSSCYCLE-YFPSCMLRTGNNFTE-SYTFEEVPR 420
DB 351 PRQSLVSTSSNABGTQITDTIGBPOLANSQFTIENTLPTLTLTGTGDEFTTGYTFNTDP 410
QY 421 FHSSYA-----HSQSLDRLMN--PLIDQYLYLNTONOGSQAONKDLF----- 463
DB 411 LKLTHTWQTNHNLACLOGITDLPISDTATSLTANAGRFSGTQONVVYVEALRTPAQ 470
QY 464 -----SFGSPAGMSVQPRKMLPGPCYRQQRVSKTKTNNNSNFMTWASKTNL 511
DB 471 IGFNQRHDFEANKGPFKVPVVP-----LDITAGSDHDAN-----GATIRNY 513
QY 512 NGR--ESTINGTA-----MASHKDDKXFFPMGSVMI FGKESAGASNTALDVM 559
DB 514 GKQHEBMAKQGAALPERYTMDALISAGRDTARCTV-----GQAPISITPPNOKOI 563
QY 560 ITDEBEIKA--TNPVATERFGTVAVNFQSSSTDPATGVDHAMGALPGWVMDRDVYLQGP 618
```

DB 564 LQREDAIAGRTMHTYNTVNSYGPUSAFPHDP-----ITPNQOI 603
QY 619 WAK---IPTDGHFHPSPLMGFGGLANPPQIILIKXTVPVNPAPAEF--SATKFSFIHQ 673
DB 604 WKKEIDLEHKRPLHATAPFV---CKNNPFGQLFVHLGF---NLTOGFDPNSTTTSRIVT- 656
QY 674 YSTGOVSVEIEMELÖKENSKRANPEVOYTSN 704
DB 657 YSTFWYKGIKFKAKLRPNLTWNPVYQATTD 687

RESULT 15

VEPCVN
coat protein VP1 - canine parvovirus (strain 780929)

N:Contains: coat protein VP2

C:Species: canine parvovirus, CPV

C>Date: 30-Jun-1987 #sequence_rev1stion 30-Jun-1987 #text_change 09-Jul-2004

C:Accession: A03702

R:Rhode III, S.L.

J. Virol. 54, 630-633, 1985

A:Title: Nucleotide sequence of the coat protein gene of canine parvovirus.

A:Reference number: A03702; NCID:85185696; PMID:3989914

A:Accession: A03702

A:Molecule type: DNA

A:Residues: 1-722 <RHO>

A:Cross-references: UNIPROT:P04863; UNIPARC:UPI0000174966; EMBL:M10989

C:Genetic8:

A:introns: 11/1

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

F:139-722/Product: coat protein VP2 #status predicted <VP2>

Query Match 9.7%; Score 385.5; DB 1; Length 722;

Best local Similarity 22.8%; Pred. No. 3, 1e-17;

Matches 178; Conservative 118; Mismatches 297; Indels 189; Gaps 35;

QY 49 GYKYLGPFGNGLDKGEVNNADAALBHDKAYDQOLKAGDNPYLRYNHADAERQERLQEDT 108
DB 11 GYKYLGPFGSLDQGEPTNPDAAKEHDEAYAYLRSKKNPYLYFSPADQRFIDQTKDAK 70
QY 109 SFGCNIGRAVFOAKKRVLEPLGLVEGAKTAPGKRPVEQSPQEPDSSSGIGTGQOPAK 168
DB 71 DMGKGIQHYFFRAKKAIAPLVLTDPHPSTSRPTK-PTKRSKPPPHIPIINLAK-----K 123
QY 169 KRANFGOTDSEVP---DPOPLGEPPA--TPAAVGPTTMASSGGA PMADNNEGADGVG 222
DB 124 KXAGAGQVQRDNIA PMSDGAVOPDGGOPAVRNERATGSGNGSGGG-----GGSGGVG 177
QY 223 NASGNMH-----CDSTWLGDRVITTTSTRTWALPTYNNHLYKQI-----SSASTGAS 268
DB 178 ISTGTNNQTEFEKFLENGWV--EITANSSRLVHLNMPESSEKDRVYVNNMDKTAIVGNMA 235
QY 269 NDNNHYFGYSTPWGYPDFNRFCHSPRDQRLINNMGFRPKRLNFKLFIQYQYEV--- 325
DB 236 LDDIHAQIVTPMSLVDANADVWFNPGDQLIIVTMSLHIVSFEQBIFFNVVLKTVSESA 295
QY 326 TNDGVTIANNLSTVQVFSDESQLPYVLGSAHQCL-----PPPPA-----DVF 371
DB 296 TQPPTKYNNNDLTASLWALDSNNTMPPTPAAMRSETLIGYPMKEFTIPTPRRYIFQWDRK 355
QY 372 MIPQYGLTLNNGSOAVGRS-----SFYCLE-YPPSOMLRGTGNNF-TFSYTEEV 419
DB 356 LISSH-----TGTSGTPTNIYHGTDPDVQFYTIENSVPVHLRTGDEFATGTFPFDC 409
QY 420 PFHSSVYHQSGLRLMPLIDQYLYLNTQONS-----GSAQNK-----D 460
DB 410 P--CRLHTWQTNBALG--LPPLINSLOSEGATNPGDIGVPODKKRGVTOGMGNTNYITE 465
QY 461 LIFRGSPAGMSVQ-----PKNMLGQPCVRQGRVSKTKTDNNNSNFTWTGASKY- 509
DB 466 ATIMRPAEVGSADYVFEASTOGPFKTLPIAAGR-----GGAQTDENQA-----ADGNRYA 517

QY 510 -NLNGRESIINPGT-----AMASHKDEDEKFEPMGVMIFGKESAGASNTALD-NVMITD 562
DB 518 FGRQHGKKTITTTGETPERFETIAHQDT-----GRYPBGDIQININPFLPVTN 564
QY 563 EEEIKATNPVATERRGTVAVNFQSSSTDPAIGDVHAMGALPGMWQDRDYLQGFIMAKI 622
DB 565 DNVLPIPIPIG-----GKTGINV--TNINFTYGLPLALNNVP-----PYVNGQIMDK 611
QY 623 PHTDGHFHPSPLMGFGGLK-----NPPQIILIKXTVPVNPAN--PPAEFSATK 666
DB 612 FDTD-----LKPRLHVNAPVCONNCPGLFVKLAPNLITNEYDPDASANMSR 658
QY 667 FASFIQYSTGOVSVEIEMELÖKENSKRANPEVOYTSNVAKSANVDFTVNNGLYTEBRP 726
DB 659 ---LVTYSDFRWKKGKLVFKAKLRASHHTNPIQOMISIN-----VDNQFNIV-P 703
QY 727 IG 728
DB 704 IG 705

Search completed: November 23, 2005, 17:44:00
Job time : 29.9904 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 23, 2005, 17:24:49 ; Search time 107.044 Seconds
(without alignments)
2458.679 Million cell updates/sec

Title: US-10-696-282-15

Perfect score: 3251
Sequence: 1 TAPGKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLTPPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003bs:*
8: geneseqp2004a:*
9: geneseqp2005a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3251	100.0	599	3 AAY71168	Aay71168 Adeno-sss
2	3251	100.0	736	3 AAY71167	Aay71167 Adeno-sss
3	3251	100.0	736	6 ABB80232	Abb80232 AAV1 vpl
4	3251	100.0	736	7 ABR62762	AbR62762 Adeno-sss
5	3251	100.0	736	7 ADE76565	AdE76565 Adeno-sss
6	3251	100.0	736	7 ADV70294	AdV70294 Primate a
7	3251	100.0	736	9 ADZ27069	AdZ27069 Adeno-sss
8	3251	100.0	736	9 ADZ27012	AdZ27012 Adeno-sss
9	3247	99.9	735	9 ADV67506	Adv67506 Amino aci
10	3233	99.4	736	9 ADZ27074	AdZ27074 Adeno-sss
11	3230	99.4	736	9 ADZ27007	AdZ27007 Adeno-sss
12	3230	99.4	736	9 ADZ27008	AdZ27008 Adeno-sss
13	3229	99.3	736	4 AAB59847	Aab59847 AAV6 caps
14	3229	99.3	736	7 ADE76566	AdE76566 Adeno-sss
15	3229	99.3	736	9 ADV70293	Adv70293 Primate a
16	3229	99.3	736	9 ADZ27070	AdZ27070 Adeno-sss
17	3203	98.5	737	9 ADZ27009	AdZ27009 Adeno-sss
18	3172.5	97.6	736	9 ADZ27086	AdZ27086 Adeno-sss
19	3163.5	97.3	737	9 ADZ27010	AdZ27010 Adeno-sss
20	2906	89.4	534	3 AAY71169	Aay71169 Adeno-sss
21	2832	87.1	736	4 AAB59846	Aab59846 AAV3B cap
22	2821.5	86.8	735	9 ADZ26996	AdZ26996 Adeno-sss
23	2820.5	86.8	735	9 ADZ26993	AdZ26993 Adeno-sss
24	2818.5	86.7	735	9 ADZ27001	AdZ27001 Adeno-sss

25	2818.5	86.7	735	9	ADZ27076	AdZ27076 Adeno-sss
26	2818.5	86.7	735	9	ADZ27000	AdZ27000 Adeno-sss
27	2818.5	86.7	735	9	ADZ27034	AdZ27034 Adeno-sss
28	2815.5	86.6	735	9	ADZ26992	AdZ26992 Adeno-sss
29	2815	86.6	736	6	AAB59845	Aab59845 AAV3A cap
30	2815	86.6	736	6	ABB80233	Abb80233 AAV3 vpl
31	2815	86.6	736	7	ABR62763	AbR62763 Adeno-sss
32	2815	86.6	736	7	ADZ27035	AdZ27035 Adeno-sss
33	2815	86.6	736	9	ADV67507	Adv67507 Amino aci
34	2815	86.6	736	9	ADZ27067	AdZ27067 Adeno-sss
35	2811.5	86.5	735	9	ADZ27002	AdZ27002 Adeno-sss
36	2808.5	86.4	735	9	ADZ26991	AdZ26991 Adeno-sss
37	2807.5	86.4	735	9	ADZ27003	AdZ27003 Adeno-sss
38	2807.5	86.4	735	9	ADZ26998	AdZ26998 Adeno-sss
39	2803.5	86.2	735	9	ADZ26994	AdZ26994 Adeno-sss
40	2803.5	86.2	735	9	ADZ27004	AdZ27004 Adeno-sss
41	2802.5	86.2	735	9	ADZ26999	AdZ26999 Adeno-sss
42	2800	86.1	734	9	ADZ27037	AdZ27037 Adeno-sss
43	2796.5	86.1	734	9	ADZ27036	AdZ27036 Adeno-sss
44	2796.5	86.0	735	9	ADZ27035	AdZ27035 Adeno-sss
45	2789.5	85.8	735	9	ADZ27006	AdZ27006 Adeno-sss

ALIGNMENTS

RESULT 1
AAY71168
ID AAY71168 strand; protein, 599 AA.
XX
AC AAY71168;
XX
DT 08-SEP-2000 (first entry)
XX
Adeno-associated virus serotype 1 capsid protein VP2.
XX
KW Adeno-associated virus serotype 1; AAV-1: rep protein; capsid protein;
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
XX vaccine; transgene; VP2.
XX
OS Adeno-associated virus 1.
XX
XX
PN MO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WC-US025694.
XX
PR 05-NOV-1998; 98US-0107114P.
XX
(UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
DR WPI, 2000-376571/32.
DR N-PSDB; AAD00772, AAD00778.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host.
XX
XX
PS Claim 7, Page 93-95; 108pp; English.
XX
XX
XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
XX which is characterised by two inverted terminal repeats (ITR) and open
XX reading frames for rep and capsid (cap) proteins. The rep reading frame
XX encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
XX reading frame encodes three structural proteins, VP1, VP2 and VP3. The
XX AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
XX regions, are useful in production of recombinant viral vectors for gene
XX delivery. These vectors can be used as gene therapy vectors, vaccine
XX vectors or antisense delivery vectors. The AAV-1 does not induce the
XX formation of neutralising antibodies specific to any serotype of AAV
XX hence is useful for transforming host cells, and in the preparation of a

CC medicament for the delivery of transgene to a host. The present sequence
CC is an AAV-1 cap protein VP2 which is useful in the production of
CC recombinant viral vector for gene delivery

XX Sequence 599 AA:

Query Match 100.0%; Score 3251; DB 3; Length 599;
Best Local Similarity 100.0%; Pred. No. 1e-253;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TAPGKKRPVEGSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDSVDPDPLGEPATPAA 60
DB 1 TAPGKKRPVEGSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDSVDPDPLGEPATPAA 60
QY 61 VGPPTMASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTNNHLY 120
DB 61 VGPPTMASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTNNHLY 120
QY 121 KOISSASTGASNDNHYFGYSTPMGYFDENRPHCHFSRPMQRLINNNMGFRPKRLNFKLF 180
DB 121 KOISSASTGASNDNHYFGYSTPMGYFDENRPHCHFSRPMQRLINNNMGFRPKRLNFKLF 180
QY 181 NIOVKEVTTNDGVTTIANNLTSTVOVPSDSEYQLPYVLGSAHQCLPPPADVPMIPQYG 240
DB 181 NIOVKEVTTNDGVTTIANNLTSTVOVPSDSEYQLPYVLGSAHQCLPPPADVPMIPQYG 240
QY 241 YLTINNGSAVGRSSFYCLEYFPSCMLRTGNNTFSYFEEVPHSSYAHQSGLDRMLNP 300
DB 241 YLTINNGSAVGRSSFYCLEYFPSCMLRTGNNTFSYFEEVPHSSYAHQSGLDRMLNP 300
QY 301 LIDQYLYLNRTQNGSAGKNDLLFSRGSAPAGMSVOPKMWLPBCYRQQRVSKTKTDNN 360
DB 301 LIDQYLYLNRTQNGSAGKNDLLFSRGSAPAGMSVOPKMWLPBCYRQQRVSKTKTDNN 360
QY 361 NSNFTWTGASKYNLNGRBSIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 420
DB 361 NSNFTWTGASKYNLNGRBSIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 420
QY 421 VMITDEEBEIKATNPVATERFGTVAVNFQSSSTDPATGVAHMGALPGMWODRDVYLQGP 480
DB 421 VMITDEEBEIKATNPVATERFGTVAVNFQSSSTDPATGVAHMGALPGMWODRDVYLQGP 480
QY 481 IMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKTTPVANPAPAFSATKFAFIQYSTG 540
DB 481 IMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKTTPVANPAPAFSATKFAFIQYSTG 540
QY 541 QVSVEIEWELQKENSKRNNPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGTRVYLTRPL 599
DB 541 QVSVEIEWELQKENSKRNNPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGTRVYLTRPL 599
```

RESULT 2

AAV71167 standard; protein; 736 AA.

AAV71167;

08-SEP-2000 (first entry)

Adeno-associated virus serotype 1 capsid protein VP1.

Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;

cap protein; recombinant viral vector; gene delivery; gene therapy;

vaccine; transgene; VPI.

Adeno-associated virus 1.

WO200028061-A2.

18-MAY-2000.

02-NOV-1999; 99WO-US025694.

PR 05-NOV-1998; 98US-0107114P.

XX (UTPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Xiao W,

XX WPI; 2000-376571/32.

DR N-PSDB; AAD00772, AAD00777.

PT Novel adeno-associated virus serotype 1 polynucleotide useful for

preparation of medicament for delivery of a transgene to a host.

PS Claim 7; Page 87-90; 108pp; English.

CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterized by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antilease delivery vectors. The AAV-1 does not induce the
CC formation of neutralizing antibodies specific to any serotype of AAV
CC hence is useful for transforming host cells, and in the preparation of a
CC medicament for the delivery of transgene to a host. The present sequence
CC is an AAV-1 cap protein VP1 which is useful in the production of
CC recombinant viral vector for gene delivery

XX Sequence 736 AA:

Query Match 100.0%; Score 3251; DB 3; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.4e-253;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TAPGKKRPVEGSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDSVDPDPLGEPATPAA 60
DB 1 TAPGKKRPVEGSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDSVDPDPLGEPATPAA 197
QY 61 VGPPTMASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTNNHLY 120
DB 61 VGPPTMASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTNNHLY 257
QY 121 KOISSASTGASNDNHYFGYSTPMGYFDENRPHCHFSRPMQRLINNNMGFRPKRLNFKLF 180
DB 121 KOISSASTGASNDNHYFGYSTPMGYFDENRPHCHFSRPMQRLINNNMGFRPKRLNFKLF 317
QY 181 NIOVKEVTTNDGVTTIANNLTSTVOVPSDSEYQLPYVLGSAHQCLPPPADVPMIPQYG 240
DB 181 NIOVKEVTTNDGVTTIANNLTSTVOVPSDSEYQLPYVLGSAHQCLPPPADVPMIPQYG 377
QY 241 YLTINNGSAVGRSSFYCLEYFPSCMLRTGNNTFSYFEEVPHSSYAHQSGLDRMLNP 300
DB 241 YLTINNGSAVGRSSFYCLEYFPSCMLRTGNNTFSYFEEVPHSSYAHQSGLDRMLNP 437
QY 301 LIDQYLYLNRTQNGSAGKNDLLFSRGSAPAGMSVOPKMWLPBCYRQQRVSKTKTDNN 360
DB 301 LIDQYLYLNRTQNGSAGKNDLLFSRGSAPAGMSVOPKMWLPBCYRQQRVSKTKTDNN 497
QY 361 NSNFTWTGASKYNLNGRBSIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 420
DB 361 NSNFTWTGASKYNLNGRBSIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 557
QY 421 VMITDEEBEIKATNPVATERFGTVAVNFQSSSTDPATGVAHMGALPGMWODRDVYLQGP 480
DB 421 VMITDEEBEIKATNPVATERFGTVAVNFQSSSTDPATGVAHMGALPGMWODRDVYLQGP 617
QY 481 IMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKTTPVANPAPAFSATKFAFIQYSTG 540
DB 481 IMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKTTPVANPAPAFSATKFAFIQYSTG 677
QY 541 QVSVEIEWELQKENSKRNNPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGTRVYLTRPL 599
DB 541 QVSVEIEWELQKENSKRNNPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGTRVYLTRPL 599
```

Db 678 QVSVEIEMELQENSKRNPEVQYTSNVAKSANVDFTVDNGLYTEPRPIGTRYLTRPL 736

RESULT 3

ABR62762

ID ABR62762 standard; protein; 736 AA.

XX ABR62762;

DT 20-NOV-2003 (first entry)

DE AAV1 vp1 protein.

KW Adeno-associated virus; AAV; serotype 9; rep; cap; vp1; vp2; vp3;

XX splice variant; transgene.

OS Adeno associated virus serotype 1.

PN WO2003052052-A2.

XX 26-JUN-2003.

XX 12-NOV-2002; 2002WO-US033631.

XX 17-DEC-2001; 2001US-0341150P.

XX 05-JUN-2002; 2002US-0386132P.

XX (UYPE-) UNIV PENNSYLVANIA.

PI Gao G, Wilson JM, Alvira M;

DR WPI; 2003-523523/49.

PT New isolated adeno-associated virus (AAV) comprising an AAV9 capsid,
useful for preparing a medicament for delivering a transgene to a cell.

PS Disclosure; Fig 2; 76pp; English.

XX The sequences given in ABR62762-34 represent vp1 proteins derived from
CC various adeno-associated virus (AAV) serotypes. These sequences were used
CC in the scope of the invention for comparison with the cap protein derived
CC from AAV serotype 9. The AAV capsid comprises three proteins vp1, vp2 and
CC vp3, which are alternative splice variants. The AAV or the nucleic acid
CC molecule is useful for preparing a medicament for delivering a transgene
CC to a cell.

XX Sequence 736 AA;

XX Query Match 100.0%; Score 3251; DB 6; Length 736;

XX Best Local Similarity 100.0%; Pred. No. 1,4e-253;

XX Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGKKRPVQSGPOBPPSSSGIGTGOQPAKKRLNFGQTGDSVDPQPLGSPPTPAA 60

DB 138 TAGKKRPPVQSGPOBPPSSSGIGTGOQPAKKRLNFGQTGDSVDPQPLGSPPTPAA 197

QY 61 VGPPTASGGGAPMADNNEGADGVNAGNMHCDSTWLGDRVITTTSTRTALPTYNHLY 120

DB 198 VGPPTASGGGAPMADNNEGADGVNAGNMHCDSTWLGDRVITTTSTRTALPTYNHLY 257

QY 121 KQISSASTGASNDNHFGYSTPMGYDFNPFHGFSPRDMQRLINNMWGFPRKLNFKLF 180

DB 258 KQISSASTGASNDNHFGYSTPMGYDFNPFHGFSPRDMQRLINNMWGFPRKLNFKLF 317

QY 181 NIQVKEVTNDGVTTANNLTSTVQVPSDEYOLPYVLGSAHOGCLPPPADVPMIPOYG 240

DB 318 NIQVKEVTNDGVTTANNLTSTVQVPSDEYOLPYVLGSAHOGCLPPPADVPMIPOYG 377

QY 241 YLTLLNNGSAVGRSSPYCLEYFSPQMLRTGNNTFFSTFEVDFPHSSYAHOSQSLDRMLNP 300

DB 378 YLTLLNNGSAVGRSSPYCLEYFSPQMLRTGNNTFFSTFEVDFPHSSYAHOSQSLDRMLNP 437

QY 301 LIDQYLYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKMLPGPCYRQORVSKTKTDNN 360

Db 438 LIDQYLYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKMLPGPCYRQORVSKTKTDNN 497

QY 361 NSNFTWGAASKYNLNGRESIINPGTAMASHKODEDKFPFPMGCVMI FGESASASTALDN 420

DB 498 NSNFTWGAASKYNLNGRESIINPGTAMASHKODEDKFPFPMGCVMI FGESASASTALDN 557

QY 421 WMITDEEIKATNPVATERFGTVAVNFOSSSTDPATGVHAMGALPGMWODRDVYLQGP 480

DB 558 WMITDEEIKATNPVATERFGTVAVNFOSSSTDPATGVHAMGALPGMWODRDVYLQGP 617

QY 481 IWAKIPHTDGHFHPSPLMGSGFGLAKNPPOILIKNTVPANPAPESATKFAFIQYSGT 540

DB 618 IWAKIPHTDGHFHPSPLMGSGFGLAKNPPOILIKNTVPANPAPESATKFAFIQYSGT 677

QY 541 QVSVEIEMELQENSKRNPEVQYTSNVAKSANVDFTVDNGLYTEPRPIGTRYLTRPL 599

DB 678 QVSVEIEMELQENSKRNPEVQYTSNVAKSANVDFTVDNGLYTEPRPIGTRYLTRPL 736

RESULT 4

ABR62762

ID ABR62762 standard; protein; 736 AA.

XX ABR62762;

DT 06-NOV-2003 (first entry)

DE Adeno associated virus 1 capsid protein vp1.

KW AAV; AAV1; capsid; vector; gene therapy; antisense therapy; vaccine.

OS Adeno associated virus.

PN WO2003052051-A2.

XX 26-JUN-2003.

XX 12-NOV-2002; 2002WO-US033630.

XX 17-DEC-2001; 2001US-0341151P.

XX 01-MAY-2002; 2002US-0377133P.

XX 05-JUN-2002; 2002US-0386122P.

XX (UYPE-) UNIV PENNSYLVANIA.

PI Gao G, Wilson JM, Alvira M;

DR WPI; 2003-523522/49.

PT New adeno-associated virus (AAV) comprising an AAV capsid having an amino
acid sequence of AAV8, useful for preparing a medicament for delivery of
a transgene to a cell and for treating cystic fibrosis or hemophilia B.

PS Disclosure; Fig 2A-C; 82pp; English.

XX The present sequence is the protein sequence of the vp1 capsid protein of
CC adeno associated virus serotype 1 (AAV1). The invention provides the
CC nucleic acid and amino acid sequences of novel AAV8 and fragments of
CC these sequences. Each of these fragments may be used in a variety of
CC vector systems and host cells. Among the desirable fragments are the cap
CC proteins, including vp1, vp2, vp3 and hypervariable regions, the rep
CC proteins, including rep78, rep68, rep52 and rep40, and the sequences
CC encoding these proteins. The fragments may be used alone, in combination
CC with other AAV8 sequences or in combination with elements from other AAV
CC or non-AAV viral sequences in the production of recombinant AAV and for
CC use as antisense delivery vectors, gene therapy vectors or vaccine
CC vectors. A claimed molecule comprises a cap protein of a functional AAV
CC rep gene from a serotype selected from AAV1, AAV2, AAV3, AAV4, AAV5 and
CC AAV6

XX Sequence 736 AA;

Db 678 QVSVEIEMLOKENS KRNPEVQYTSNYSKANSANVDTVNNGLYTERPRIGRIYLRLPL 736

RESULT 6
ADV70294
ADV70294 standard; protein, 736 AA.

ADV70294;
10-MAR-2005 (first entry)

Primate adeno-associated virus 1 capsid protein VP1.

immunosuppressive; gene therapy; immunogenicity; virus inactivation;
hemophilia; pepck deficiency; galactosemia; phenylketonuria;
Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
achroscleorosis; thrombosis; embolism; Parkinson's disease;
congestive heart failure; cancer; inflammation; immune disorder;
muscular dystrophy; diabetes; VPI.

Adeno-associated virus 1.

MO2004112727-A2.

29-DEC-2004.

21-JUN-2004; 2004MO-US019884.

19-JUN-2003; 2003US-0480395P.

30-APR-2004; 2004US-0567310P.

03-JUN-2004; 2004US-0576501P.

(AVIG-) AVIGEN INC.

Arbetman AE, Colosi P, Lochrie MA, Surosky RT;

WPI; 2005-048755/05.

New mutated adeno-associated virus (AAV) capsid protein that when present in an AAV virion imparts decreased immunoreactivity to the virion as compared to the corresponding wild-type virion, useful for treating e.g. hemophilia.

Example 5; SEQ ID NO 20; 136pp; English.

The invention describes a mutated adeno-associated virus (AAV) capsid

protein that when present in an AAV virion imparts decreased

immunoreactivity to the virion as compared to the corresponding wild-type

virion. Also described are: a polynucleotide encoding the mutated protein

above; a recombinant AAV virion comprising the mutated protein above; and

delivering a recombinant AAV virion to a cell or tissue of a vertebrate

subject. The recombinant AAV virion is useful for delivering a

heterologous nucleic acid molecule to a cell or tissue of a vertebrate

subject, where the protein encoded by the heterologous nucleic acid

molecule is expressed at a level that provides a therapeutic effect,

where the recombinant AAV virion may comprise a non-primate, mammalian

AAV capsid protein that when present in an AAV virion imparts decreased

immunoreactivity to the virion as compared to immunoreactivity of primate

AAV-2, and the heterologous nucleic acid molecule, where the heterologous

nucleic acid molecule encodes a therapeutic protein and is operably

linked to control elements capable of directing the in vivo transcripion

and translation of the protein. The protein or the recombinant AAV virion

is useful for treating or preventing a wide variety of disorders such as

hemophilia, glycogen storage deficiency type 1A, pepck deficiency,

galactosemia, phenylketonuria, Crigler-Najjar disease, Gout and Lesch-

Nyan syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,

thrombosis, embolisms, Parkinson's disease, congestive heart failure,

cancer, inflammatory and immune disorders, muscular dystrophies, and

diabetes. This is the amino acid sequence of adeno-associated virus 1

(AAV1) capsid protein VP1.

Sequence 736 AA;

Query Match 100.0%; Score 3251; DB 9; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.4e-253;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TAPGKKRPVEGSPQPDSSSGIGKTGOQPAKKRLNFGQTGDBSEYVDPDQPLGEPPATPA 60

138 TAPGKKRPVEGSPQPDSSSGIGKTGOQPAKKRLNFGQTGDBSEYVDPDQPLGEPPATPA 197

61 VGPPTMASGGGAPMADNNEGADVGNAAGNMHCDSITWLGDRVITTSITRTWALPTYNHLY 120

198 VGPPTMASGGGAPMADNNEGADVGNAAGNMHCDSITWLGDRVITTSITRTWALPTYNHLY 257

121 KOISASISGASNDNHYFGSTPMGYPDFRRFHCHSPSPDMQMLINNNGFFRKRLNFKLP 180

258 KOISASISGASNDNHYFGSTPMGYPDFRRFHCHSPSPDMQMLINNNGFFRKRLNFKLP 317

181 NIQVEKVTINDGVTTIANNLSTVQVFSDBSEYQLPVLGSAHQGLPPPADVFMIPOYG 240

318 NIQVEKVTINDGVTTIANNLSTVQVFSDBSEYQLPVLGSAHQGLPPPADVFMIPOYG 377

241 YLTINNNGSQAAGRSSFYCLEYFPPSQMLRTGNNFTSYEEVPHSSYAHSGSLDRLAMP 300

378 YLTINNNGSQAAGRSSFYCLEYFPPSQMLRTGNNFTSYEEVPHSSYAHSGSLDRLAMP 437

301 LIDQYLYINRTONSGSAQNKDLLFSRGSPPAGMSVQPPKWLPGPCYRQQRVSKTTDNN 360

438 LIDQYLYINRTONSGSAQNKDLLFSRGSPPAGMSVQPPKWLPGPCYRQQRVSKTTDNN 497

361 NSNFTMTGASKYNLNGRESIINRGTAASHKXDEDFPFMSGVMIFGKSAGASNTALDN 420

498 NSNFTMTGASKYNLNGRESIINRGTAASHKXDEDFPFMSGVMIFGKSAGASNTALDN 557

421 WMITDEEELKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRVYLOGP 480

558 WMITDEEELKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRVYLOGP 617

481 IMAKIPHTDGHHPSPFLMGFGFKNPPOILLKNTVPANPPAEFATPFASITTOYSTG 540

618 IMAKIPHTDGHHPSPFLMGFGFKNPPOILLKNTVPANPPAEFATPFASITTOYSTG 677

541 QVSVEIEMLOKENS KRNPEVQYTSNYSKANSANVDTVNNGLYTERPRIGRIYLRLPL 599

678 QVSVEIEMLOKENS KRNPEVQYTSNYSKANSANVDTVNNGLYTERPRIGRIYLRLPL 736

RESULT 7

AD227069
ID AD227069 standard; protein, 736 AA.

AD227069;
30-JUN-2005 (first entry)

Adeno-associated virus protein SEQ ID NO 219.

rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
Kw scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
Kw bacterial infection; cancer; ulcerative colitis; antineoplastic;
Kw antirheumatic; neuroprotective; antiinflammatory; antidiabetic;
Kw antipsychotic; vasotropic; gastrointestinal gen.; hemostatic; anti-HIV;
Kw antiviral; antibacterial; cyostatic; antiviral; dermatological.

Adeno-associated virus.

MO200503321-A2.

14-APR-2005.

30-SEP-2004; 2004MO-US028817.

30-SEP-2003; 2003US-0508226P.

29-APR-2004; 2004US-0566546P.

```

XX (UYPE-) UNIV PENNSYLVANIA.
PA
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
XX WPI, 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
XX Discloure; SEQ ID NO 219; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
XX Sequence 736 AA:
SQ
Query Match 100.0%; Score 3251; DB 9; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.4e-253;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAPGKKRPVEGSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSVDPDQPLGEPATPAA 60
DB 138 TAPGKKRPVEGSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSVDPDQPLGEPATPAA 197
QY 61 VGPTTMAAGGAPMADNNEGADGVGNASGNHCDSTWLGADRVITTTSTWMLPTNNHLY 120
DB 198 VGPTTMAAGGAPMADNNEGADGVGNASGNHCDSTWLGADRVITTTSTWMLPTNNHLY 257
QY 121 KOISSASTGASNDNHYPGYSTPMGYFDNRFHCHFSPRDMQRLINNMGFRPKRLNFKLF 180
DB 258 KOISSASTGASNDNHYPGYSTPMGYFDNRFHCHFSPRDMQRLINNMGFRPKRLNFKLF 317
QY 181 NIQVEVTTNDGVTTIANNLSTVQVFSDSSEYOLPYVIGSAHQGLPPPADVPMIPOYG 240
DB 318 NIQVEVTTNDGVTTIANNLSTVQVFSDSSEYOLPYVIGSAHQGLPPPADVPMIPOYG 377
QY 241 YTLTNGSQAQVGRSSPYCLEYFPSCQMLRTGNNFTSEYEEVPHSSYAHOSQSLDRLAMP 300
DB 378 YTLTNGSQAQVGRSSPYCLEYFPSCQMLRTGNNFTSEYEEVPHSSYAHOSQSLDRLAMP 437
QY 301 LIDQYLYVINRTQNGSAQNKDLFSGSGPAGMSVQPPNMLPGCYROQRVSKTKTDNN 360
DB 438 LIDQYLYVINRTQNGSAQNKDLFSGSGPAGMSVQPPNMLPGCYROQRVSKTKTDNN 497
QY 361 NSNFTWTGASKYNLNRESIINPGTAMASHKXDEDEDFPMGSMVMTFGKSSAASNTALDN 420
DB 498 NSNFTWTGASKYNLNRESIINPGTAMASHKXDEDEDFPMGSMVMTFGKSSAASNTALDN 557
QY 421 WMTDEEBIKATNPVATERFGTVAVNFQSSSTDPAQDVHAMGALPGMWQDRDYLQOP 480
DB 558 WMTDEEBIKATNPVATERFGTVAVNFQSSSTDPAQDVHAMGALPGMWQDRDYLQOP 617
QY 481 IMAKIPHTDGHFHPSPFLMGFGKLPKPPQILIKNTVPANPPAEFSATKFAFITQYSTG 540
DB 618 IMAKIPHTDGHFHPSPFLMGFGKLPKPPQILIKNTVPANPPAEFSATKFAFITQYSTG 677
QY 541 QVSVIEBMLQKENSRRANPEVOYTSYNAKSAVNDPTVNNGLYTPERRIGRYYLTRPL 599
DB 678 QVSVIEBMLQKENSRRANPEVOYTSYNAKSAVNDPTVNNGLYTPERRIGRYYLTRPL 736

```

```

RESULT 8
AD227012
XX ID AD227012 standard; protein: 736 AA.
XX
XX AC AD227012;
XX
XX 30-JUN-2005 (first entry)
XX
XX DE Adeno-associated virus protein SEQ ID NO 162.
XX
XX Rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
XX antiparrotic; vasotropic; gastrointestinal-gen; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
XX Adeno-associated virus.
XX
XX MO200503321-A2.
XX
XX 14-APR-2005.
XX
XX 30-SEP-2004; 2004WO-US028917.
XX
XX 30-SEP-2003; 2003US-0508226P.
XX
XX 29-APR-2004; 2004US-056546P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
XX WPI, 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
XX Discloure; SEQ ID NO 162; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
XX Sequence 736 AA:
SQ
Query Match 100.0%; Score 3251; DB 9; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.4e-253;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAPGKKRPVEGSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSVDPDQPLGEPATPAA 60
DB 138 TAPGKKRPVEGSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSVDPDQPLGEPATPAA 197
QY 61 VGPTTMAAGGAPMADNNEGADGVGNASGNHCDSTWLGADRVITTTSTWMLPTNNHLY 120
DB 198 VGPTTMAAGGAPMADNNEGADGVGNASGNHCDSTWLGADRVITTTSTWMLPTNNHLY 257
QY 121 KOISSASTGASNDNHYPGYSTPMGYFDNRFHCHFSPRDMQRLINNMGFRPKRLNFKLF 180
DB 258 KOISSASTGASNDNHYPGYSTPMGYFDNRFHCHFSPRDMQRLINNMGFRPKRLNFKLF 317
QY 181 NIQVEVTTNDGVTTIANNLSTVQVFSDSSEYOLPYVIGSAHQGLPPPADVPMIPOYG 240

```

```

Db      318 NIQKEVTINDGVTTIANNLSTVOVFSDEYQLPYVLGSAHQGLPPPADVFMIPOYG 377
Qy      241 YLTNNGSAVGRSSPYCLEYFPSSQMLRTGNFTFSYTEEVPFHSSVAHSQSLDLRLMP 300
Db      378 YLTNNGSAVGRSSPYCLEYFPSSQMLRTGNFTFSYTEEVPFHSSVAHSQSLDLRLMP 437
Qy      301 LIDQYLYLNRTQNGSGAQNKDLPFRSGSPAGMSVQPKWMLPGPCYRQORVSKTDDNN 360
Db      438 LIDQYLYLNRTQNGSGAQNKDLPFRSGSPAGMSVQPKWMLPGPCYRQORVSKTDDNN 497
Qy      361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMITGKSAGASNTALDN 420
Db      498 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMITGKSAGASNTALDN 557
Qy      421 WMTDEEIKATNPVATERFGTVAVNFQSSSTDPAVDHAMGALPGMWQDRDYLQGP 480
Db      558 WMTDEEIKATNPVATERFGTVAVNFQSSSTDPAVDHAMGALPGMWQDRDYLQGP 617
Qy      481 IWAKIPHTDGHFHPSPLMGFGFKNPPOILLIKNTVPANPAEFSATYFASFTIYSTG 540
Db      618 IWAKIPHTDGHFHPSPLMGFGFKNPPOILLIKNTVPANPAEFSATYFASFTIYSTG 677
Qy      541 QVSVEIEMELQENSKRNPNPEVOYTSNYAKSANVDFTVNNGLYTPRPRTGRYLTRPL 599
Db      678 QVSVEIEMELQENSKRNPNPEVOYTSNYAKSANVDFTVNNGLYTPRPRTGRYLTRPL 736

RESULT 9
ADV67506 standard; protein; 735 AA.
ID      ADV67506
XX      XX
AC      ADV67506;
XX      XX
DT      10-MAR-2005 (first entry)
XX      XX
DE      Amino acid sequence of the capsid protein of AAV serotype 1.
XX      XX
KM      antiarteriosclerotic; antilipemic; gene therapy; cholesterol;
KM      apolipoprotein E; apob; apolipoprotein A; apoa; atherosclerosis;
KM      lipoprotein defect; capsid protein.
XX      XX
OS      Adeno-associated virus.
XX      XX
PN      W02004108922-A2.
XX      XX
PD      16-DEC-2004.
XX      XX
PF      23-APR-2004; 2004WO-US010965.
XX      XX
PR      25-APR-2003; 2003US-0465293P.
XX      XX
PA      (UYPR-) UNIV PENNSYLVANIA.
XX      XX
PI      Rader DJ, Wilson JM;
XX      XX
DR      WPI; 2005-031700/03.
XX      XX
PT      Lowering total cholesterol levels and treating atherosclerosis in a
PT      subject comprises delivering a recombinant adeno-associated virus (AAV)
PT      comprising an AAV serotype capsid protein or a gene encoding human
PT      apolipoprotein E (apoE) or apoA.
XX      XX
PS      Disclosure; SEQ ID NO 4; 69pp; English.
XX      XX
CC      The specification describes a method for lowering total cholesterol
CC      levels in a subject. The method comprises delivering to the subject a
CC      recombinant adeno-associated virus (AAV) comprising a gene encoding a
CC      human apolipoprotein E (apoE) or apoA under the control of a regulatory
CC      control sequence which directs expression of the gene. The recombinant
CC      AAV also comprises a capsid protein selected from an AAV serotype (e.g.,
CC      AAV7 or AAV8) which preferentially expresses high levels of transgene in
CC      liver. A therapeutically effective amount of apoE or apoA expression is
CC      obtained upon delivery of low dose of AAV. The method of the invention is

```

CC useful for lowering total cholesterol levels in a subject, e.g. for
 CC treating atherosclerosis, and for correcting defects in lipoprotein. The
 CC present sequence represents a capsid protein of AAV serotype 1 (AAV1),
 CC which may be used in recombinant AAV vectors of the invention.

Seq Sequence 735 AA;

Query Match 99.9%; Score 3247; DB 9; Length 735;
 Best Local Similarity 100.0%; Pred. No. 2,9e-253;
 Matches 598; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 TAPGKKRPVEGSPQEPDSSGIGKTGOQPAKKRLNFGQDSESYDDPQPLGEPPATPA 60
Db      138 TAPGKKRPVEGSPQEPDSSGIGKTGOQPAKKRLNFGQDSESYDDPQPLGEPPATPA 197
Qy      61 VGPPTMASGGGAPMADNNEGADGVNAGSNMHCSTWLGDRIYTTSTRTWALPTNNHLY 120
Db      198 VGPPTMASGGGAPMADNNEGADGVNAGSNMHCSTWLGDRIYTTSTRTWALPTNNHLY 257
Qy      121 KOISSASTGASNDNHYFGYSTWGYFDPRFHCHSPRPMORLINNMGRPRRLNPKLF 180
Db      258 KOISSASTGASNDNHYFGYSTWGYFDPRFHCHSPRPMORLINNMGRPRRLNPKLF 317
Qy      181 NIQKEVTINDGVTTIANNLSTVOVFSDEYQLPYVLGSAHQGLPPPADVFMIPOYG 240
Db      318 NIQKEVTINDGVTTIANNLSTVOVFSDEYQLPYVLGSAHQGLPPPADVFMIPOYG 377
Qy      241 YLTNNGSAVGRSSPYCLEYFPSSQMLRTGNFTFSYTEEVPFHSSVAHSQSLDLRLMP 300
Db      378 YLTNNGSAVGRSSPYCLEYFPSSQMLRTGNFTFSYTEEVPFHSSVAHSQSLDLRLMP 437
Qy      301 LIDQYLYLNRTQNGSGAQNKDLPFRSGSPAGMSVQPKWMLPGPCYRQORVSKTDDNN 360
Db      438 LIDQYLYLNRTQNGSGAQNKDLPFRSGSPAGMSVQPKWMLPGPCYRQORVSKTDDNN 497
Qy      361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMITGKSAGASNTALDN 420
Db      498 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMITGKSAGASNTALDN 557
Qy      421 WMTDEEIKATNPVATERFGTVAVNFQSSSTDPAVDHAMGALPGMWQDRDYLQGP 480
Db      558 WMTDEEIKATNPVATERFGTVAVNFQSSSTDPAVDHAMGALPGMWQDRDYLQGP 617
Qy      481 IWAKIPHTDGHFHPSPLMGFGFKNPPOILLIKNTVPANPAEFSATYFASFTIYSTG 540
Db      618 IWAKIPHTDGHFHPSPLMGFGFKNPPOILLIKNTVPANPAEFSATYFASFTIYSTG 677
Qy      541 QVSVEIEMELQENSKRNPNPEVOYTSNYAKSANVDFTVNNGLYTPRPRTGRYLTRP 598
Db      678 QVSVEIEMELQENSKRNPNPEVOYTSNYAKSANVDFTVNNGLYTPRPRTGRYLTRP 735

RESULT 10
ADZ27074 standard; protein; 736 AA.
ID      ADZ27074
XX      XX
AC      ADZ27074;
XX      XX
DT      30-JUN-2005 (first entry)
XX      XX
DE      Adeno-associated virus protein SEQ ID NO 224.
XX      XX
KM      rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM      scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM      bacterial infection; cancer; ulcerative colitis; antineoplastic;
KM      antihistaminic; neuroprotective; antiinflammatory; antidiabetic;
KM      antiparasitic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM      virucide; antibacterial; cytostatic; anticancer; dermatological.
XX      XX
OS      Adeno-associated virus.
XX      XX
PN      W02005033321-A2.

```

PD 14-APR-2005.
 XX 30-SEP-2004; 2004MO-US028817.
 XX 30-SEP-2003; 2003US-0508226P.
 PR 29-APR-2004; 2004US-0566546P.
 XX (UTYPE-) UNIV PENNSYLVANIA.
 PA
 XX Wilson JM, Gao G, Alvirra MR, Vandenberghe LH;
 PI MPI, 2005-285437/29.
 XX
 PT New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.
 XX
 XX Claim 19; SEQ ID NO 224; 569pp; English.
 PS
 CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.
 XX
 XX Sequence 736 AA;
 SQ

Query Match 99.4%; Score 3233; DB 9; Length 736;
 Best Local Similarity 99.5%; Pred. No. 3.9e-252;
 Matches 596; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGSESVDPDQPGEPATPAA 60
 DB 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGSESVDPDQPGEPATPAA 197
 QY 61 VGPPTMASGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTMALPTNNHLY 120
 DB 198 VGPPTMASGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTMALPTNNHLY 257
 QY 121 KOISSASTGASNDNHFGYSTPWGTFDENRFCHFSPRDMORLNNNGFRPKRLNFKLF 180
 DB 258 KOISSASTGASNDNHFGYSTPWGTFDENRFCHFSPRDMORLNNNGFRPKRLNFKLF 317
 QY 181 NIOVEVTTNDGVTTIANNLSTVOVFSDEYQLPVYLSAQGCLPPRPADVPMIPOYG 240
 DB 318 NIOVEVTTNDGVTTIANNLSTVOVFSDEYQLPVYLSAQGCLPPRPADVPMIPOYG 377
 QY 241 YTLNNGSAQVGRSSFFCLEYFSPQMLRTGNNFTFSYTFEEVYFHSYAHOSGLRLMNP 300
 DB 378 YTLNNGSAQVGRSSFFCLEYFSPQMLRTGNNFTFSYTFEEVYFHSYAHOSGLRLMNP 437
 QY 301 LIDQYLYYNRTONOGSAQONKDLFSRGS PAKMSVQPRNWLPGPCYRQQRVSKTKTDNN 360
 DB 438 LIDQYLYYNRTONOGSAQONKDLFSRGS PAKMSVQPRNWLPGPCYRQQRVSKTKTDNN 497
 QY 361 NNNFTMTGASKYNLNRRESITINPGTAMASHKXDEDEFFPMGQVMIFQKXSASASTALDN 420
 DB 498 NNNFTMTGASKYNLNRRESITINPGTAMASHKXDEDEFFPMGQVMIFQKXSASASTALDN 557
 QY 421 WATDEEBIKATNPVATERFGTVAANFQSSSTDPATGDVHAMGALPGMWQDRDYLQGP 480
 DB 558 WATDEEBIKATNPVATERFGTVAANFQSSSTDPATGDVHAMGALPGMWQDRDYLQGP 617
 QY 481 IAKIPHTDGHFHPSPFLMGFGILKNPPQILIKNTFVPANPPAFSATKFAFTIOYSTG 540
 DB 618 IAKIPHTDGHFHPSPFLMGFGILKNPPQILIKNTFVPANPPAFSATKFAFTIOYSTG 677

QY 541 QVSVEIEMELQKENS KKMNPVEQYTSNYSKANSANDFTVDNNGLYTERPRIGTRLYTTPPL 599
 DB 678 QVSVEIEMELQKENS KKMNPVEQYTSNYSKANSANDFTVDNNGLYTERPRIGTRLYTTPPL 736

RESULT 11
 ADZ27007
 ID ADZ27007 standard; protein; 736 AA.
 AC
 AC ADZ27007;
 DT
 DT 30-JUN-2005 (first entry)
 XX
 DE Adeno-associated virus protein SEQ ID NO 157.
 XX

KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KM bacterial infection; cancer; ulcerative colitis; antirheumatic;
 KM antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
 KM antiproliferative; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 KM virucide; antibacterial; cytostatic; antiulcer; dermatological.
 XX
 OS Adeno-associated virus.
 XX
 PD WO200503321-A2.
 XX
 PN 14-APR-2005.
 XX
 PF 30-SEP-2004; 2004MO-US028817.
 XX
 PR 30-SEP-2003; 2003US-0508226P.
 XX
 PR 29-APR-2004; 2004US-0566546P.
 XX

(UTYPE-) UNIV PENNSYLVANIA.

PI Wilson JM, Gao G, Alvirra MR, Vandenberghe LH;
 PI MPI, 2005-285437/29.
 XX
 DR

PT New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.
 XX

PS Claim 19; SEQ ID NO 157; 569pp; English.

CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.
 XX
 XX Sequence 736 AA;
 SQ

Query Match 99.4%; Score 3230; DB 9; Length 736;
 Best Local Similarity 99.2%; Pred. No. 6.9e-252;
 Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGSESVDPDQPGEPATPAA 60
 DB 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGSESVDPDQPGEPATPAA 197
 QY 61 VGPPTMASGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTMALPTNNHLY 120
 DB 198 VGPPTMASGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTMALPTNNHLY 257

```

QY 121 KOISSASTGASNDNHFGYSTWGYFDNRFHCHFSPPRMOILLNNMGFRKRLNFKLF 180
DB 258 KOISSTSTGASNDNHFGYSTWGYFDNRFHCHFSPPRMOILLNNMGFRKRLNFKLF 317
QY 181 NIQVEVTTNDGVTTIANNLSTVOFSDSEYQLPYVLGSAHQGLPPPAVPMIPOYG 240
DB 318 NIQVEVTTNDGVTTIANNLSTVOFSDSEYQLPYVLGSAHQGLPPPAVPMIPOYG 377
QY 241 YTLNNGSQA VGRSSFCLEFPSCQMLRTGNFTFSYTFEEVPHSSVYAHSGSLDLRLMP 300
DB 378 YTLNNGSQA VGRSSFCLEFPSCQMLRTGNFTFSYTFEEVPHSSVYAHSGSLDLRLMP 437
QY 301 LIDQYLYLNRTONOGSAGNOKDLFSRGS PAMGSVQPKNWLPGPCYRQORVSKTTDNN 360
DB 438 LIDQYLYLNRTONOGSAGNOKDLFSRGS PAMGSVQPKNWLPGPCYRQORVSKTTDNN 497
QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDEFFPMGSGVMTFGKESAGASNTALDN 420
DB 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDEFFPMGSGVMTFGKESAGASNTALDN 557
QY 421 VMTDEEELKATNPVATERFGTVA VNFQSSSTDPRATGDVHAMGALPGMWQDRDYLQGP 480
DB 558 VMTDEEELKATNPVATERFGTVA VNFQSSSTDPRATGDVHAMGALPGMWQDRDYLQGP 617
QY 481 IWAKIPHTDGHFHPSPFLMGFGGLKNPPOILLKNTVPANPAEFSATYFASFTIYOSTG 540
DB 618 IWAKIPHTDGHFHPSPFLMGFGGLKNPPOILLKNTVPANPAEFSATYFASFTIYOSTG 677
QY 541 QVSVEIEMELQKENS KRNMPDEVQYTSNYAKSANVDFTVNNGLYTEPRPIGRYLRPL 599
DB 678 QVSVEIEMELQKENS KRNMPDEVQYTSNYAKSANVDFTVNNGLYTEPRPIGRYLRPL 736

```

RESULT 12

AD227008 standard; protein; 736 AA.

AD227008;

30-JUN-2005 (first entry)

Adeno-associated virus protein SEQ ID NO 158.

KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KM bacterial infection; cancer; ulcerative colitis; antineoplastic;
 KM antirheumatic; neutroprotective; antiinflammatory; antidiabetic;
 KM antiparasitic; vasodilator; gastrointestinal-gen.; hemostatic; anti-HIV;
 KM virucide; antibacterial; cytostatic; anticancer; dermatological.

Adeno-associated virus.

W0200503321-A2.

14-APR-2005.

30-SEP-2004; 2004MO-US028817.

30-SEP-2003; 2003JUS-0508226P.

29-APR-2004; 2004US-056546P.

(UYPR-) UNIV PENNSYLVANIA.

W11son UM, Gao G, Alvira MR, Vandenberghe LH;

WPI; 2005-285437/29.

PT New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.

Claim 19; SEQ ID NO 158; 569pp; English.

XX The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.

SQ Sequence 736 AA;

Query Match 99.4%; Score 3230; DB 9; Length 736;

Best Local Similarity 99.5%; Pred. No. 6; 9e-252; Indels 0; Gaps 0;

Matches 596; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 TAPGKKRVEQSGPOBPDSSGIGKTGOOPAKKRLNFGOTGSESVDPDQPLGEPPATPA 60
DB 138 TAPGKKRVEQSGPOBPDSSGIGKTGOOPAKKRLNFGOTGSESVDPDQPLGEPPATPA 197
QY 61 VGPPTMAGGGA PMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTWALPTYNNHLY 120
DB 198 VGPPTMAGGGA PMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTWALPTYNNHLY 257
QY 121 KOISSASTGASNDNHFGYSTWGYFDNRFHCHFSPPRMOILLNNMGFRKRLNFKLF 180
DB 258 KOISSASTGASNDNHFGYSTWGYFDNRFHCHFSPPRMOILLNNMGFRKRLNFKLF 317
QY 181 NIQVEVTTNDGVTTIANNLSTVOFSDSEYQLPYVLGSAHQGLPPPAVPMIPOYG 240
DB 318 NIQVEVTTNDGVTTIANNLSTVOFSDSEYQLPYVLGSAHQGLPPPAVPMIPOYG 377
QY 241 YTLNNGSQA VGRSSFCLEFPSCQMLRTGNFTFSYTFEEVPHSSVYAHSGSLDLRLMP 300
DB 378 YTLNNGSQA VGRSSFCLEFPSCQMLRTGNFTFSYTFEEVPHSSVYAHSGSLDLRLMP 437
QY 301 LIDQYLYLNRTONOGSAGNOKDLFSRGS PAMGSVQPKNWLPGPCYRQORVSKTTDNN 360
DB 438 LIDQYLYLNRTONOGSAGNOKDLFSRGS PAMGSVQPKNWLPGPCYRQORVSKTTDNN 497
QY 421 VMTDEEELKATNPVATERFGTVA VNFQSSSTDPRATGDVHAMGALPGMWQDRDYLQGP 480
DB 558 VMTDEEELKATNPVATERFGTVA VNFQSSSTDPRATGDVHAMGALPGMWQDRDYLQGP 617
QY 481 IWAKIPHTDGHFHPSPFLMGFGGLKNPPOILLKNTVPANPAEFSATYFASFTIYOSTG 540
DB 618 IWAKIPHTDGHFHPSPFLMGFGGLKNPPOILLKNTVPANPAEFSATYFASFTIYOSTG 677
QY 541 QVSVEIEMELQKENS KRNMPDEVQYTSNYAKSANVDFTVNNGLYTEPRPIGRYLRPL 599
DB 678 QVSVEIEMELQKENS KRNMPDEVQYTSNYAKSANVDFTVNNGLYTEPRPIGRYLRPL 736

```

RESULT 13

AAB59847 standard; protein; 736 AA.

AAB59847;

28-MAR-2001 (first entry)

AAV6 capsid protein VP1.

KM AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
 KM atherosclerosis; sickle cell anemia; thalassemia;
 KM blood clotting disorder; diabetes; capsid protein VP1.

XX Adeno associated virus.
 OS US6156303-A.
 PN 05-DEC-2000.
 PD 11-JUN-1997; 97US-00873168.
 PF 11-JUN-1997; 97US-00873168.
 PR 11-JUN-1997; 97US-00873168.
 XX (UNIM) UNIV WASHINGTON.
 PA Russell DW, Rutledge EA;
 PI WPI; 2001-060164/07.
 DR Adeno-associated virus serotype 6 and viral vector derived from it for
 XX gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
 PT syndrome, sickle cell anemia, thalassemia and diabetes.
 PR
 XX Disclosure; Fig 2; 50pp; English.
 PS
 XX The present invention relates to adeno-associated virus serotypes. The
 CC present sequence is capsid protein vp1 of one such serotype (AAV6). AAV6
 CC can be used to construct AAV viral vectors for use in gene therapy for a
 CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
 CC sickle cell anaemia, thalassemia, blood clotting disorders and diabetes.
 CC The AAV viral vectors have increased transduction efficiency of a
 CC particular host cell as the AAV virion containing the AAV vector genome
 CC can be modified to express a capsid protein of an AAV serotype that
 CC transduces the selected host cell
 CC
 SO Sequence 736 AA;

Query Match 99.3%; Score 3229; DB 4; Length 736;
 Best Local Similarity 99.2%; Pred. No. 8.3e-252;
 Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAPGKKRPVPOBPDPSSSGIGTKGQOPAKKRLNFGQTGSSEVPDPOLGEBPATPAA 60
 DB 138 TAPGKKRPVPOBPDPSSSGIGTKGQOPAKKRLNFGQTGSSEVPDPOLGEBPATPAA 197
 QY 61 VGPPTMASGGAPMADNNGAGDVGAGNAGNHCDSWTGLDRVITTTSTRTWALPTNNHLY 120
 DB 138 VGPPTMASGGAPMADNNGAGDVGAGNAGNHCDSWTGLDRVITTTSTRTWALPTNNHLY 257
 QY 121 KQISSASTGASNDNHFGYSTPMGYFDENRFCHFSPPRDMORLNNNGFPRKRLNFKLF 180
 DB 258 KQISSASTGASNDNHFGYSTPMGYFDENRFCHFSPPRDMORLNNNGFPRKRLNFKLF 317
 QY 181 NIQVEKVTNNDOVTITANNLTSTVQVFSSEVQLPVYLGASAOGLCPRPAPVFMIPQYG 240
 DB 318 NIQVEKVTNNDOVTITANNLTSTVQVFSSEVQLPVYLGASAOGLCPRPAPVFMIPQYG 377
 QY 241 YTTLNGSOAVGRSSPYCLEYFPSSQMLRTGNNTFSYEEVFPFSSYAHSSQSLRLNMP 300
 DB 378 YTTLNGSOAVGRSSPYCLEYFPSSQMLRTGNNTFSYEEVFPFSSYAHSSQSLRLNMP 437
 QY 301 LIDQYLTYLNRTOGSAQNKDLLFSRGSPPAGMSVQPRNMLPGPCYRQORVSKTKTDNN 360
 DB 438 LIDQYLTYLNRTOGSAQNKDLLFSRGSPPAGMSVQPRNMLPGPCYRQORVSKTKTDNN 497
 QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDEKFFPMGSVMIPGKSSASASTALDN 420
 DB 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDEKFFPMGSVMIPGKSSASASTALDN 557
 QY 421 WITDEEBEIKATNPVATERRGTVAANVFOSSSTDPTAGDVHANGALPGMWQORDVYLOGP 480
 DB 558 WITDEEBEIKATNPVATERRGTVAANVFOSSSTDPTAGDVHANGALPGMWQORDVYLOGP 617
 QY 481 IWAKIPHTDGHFHPSPLMGFGGLKNPPOILLKNTVPANPAPFASATKPSFITOYSTG 540

DB 618 IWAKIPHTDGHFHPSPLMGFGGLKNPPOILLKNTVPANPAPFASATKPSFITOYSTG 677
 QY 541 QVSVEIEMELQKENSKRKNPEVOYTSNYAKSANDVFTVDNNGLTERTPRPIGTRILYTRPL 599
 DB 678 QVSVEIEMELQKENSKRKNPEVOYTSNYAKSANDVFTVDNNGLTERTPRPIGTRILYTRPL 736

RESULT 14
 ADE76566
 ID ADE76566 standard; protein; 736 AA.
 AC ADE76566;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX

Adeno-associated virus (AAV) related protein, SEQ ID No 65.

XX adeno-associated virus; AAV; cytostatic; antiproliferative; antirheumatic;
 XX antiarthritic; neuroprotective; antidiabetic; antithyroid;
 KW dermatological; antiinflammatory; gene therapy; vaccine;
 KW hyperproliferative; cancer; psoriasis; autoimmune disease;
 KW rheumatoid arthritis; multiple sclerosis; diabetes;
 KW autoimmune thyroiditis; scleroderma; Crohn's disease.
 KW
 XX

Unidentified.

EP310571-A2.

14-MAY-2003.

12-NOV-2002; 2002EP-00257826.

13-NOV-2001; 2001US-0350607P.

17-DEC-2001; 2001US-0341117P.

01-MAY-2002; 2002US-0377066P.

05-JUN-2002; 2002US-0386675P.

(UNPE-) UNIV PENNSYLVANIA.

Gao G, Wilson JM, Alvira M;

WPI; 2003-450984/43.

Detecting adeno-associated virus sequences in a sample, useful for e.g.

preventing or treating hyperproliferative or autoimmune diseases,

comprising subjecting a sample having a DNA to amplification via

polymerase chain reaction.

Disclosure; SEQ ID NO 65; 419pp; English.

The invention relates to a novel method for detecting adeno-associated virus (AAV) sequences in a sample, which comprises subjecting a sample containing a DNA to amplification via a polymerase chain reaction (PCR). The AAV sequence have the following activities: cytostatic, antiproliferative, antirheumatic, antiarthritic, neuroprotective, antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV sequence can be used in gene therapy or as part of a vaccine to treat disorders. The method is useful in detecting and/or identifying AAV sequences and isolating novel sequences that are identified. The AAV sequences may be used e.g. for preventing or treating hyperproliferative conditions such as cancers and psoriasis, and other autoimmune diseases like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune thyroiditis, scleroderma or Crohn's disease. This sequence represents an AAV related protein sequence of the invention.

Sequence 736 AA;

Query Match 99.3%; Score 3229; DB 7; Length 736;
 Best Local Similarity 99.2%; Pred. No. 8.3e-252;
 Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAPGKKRPVPOBPDPSSSGIGTKGQOPAKKRLNFGQTGSSEVPDPOLGEBPATPAA 60

DB 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQDSESVDPQPLGEPATPPA 197
QY 61 VGPPTWASGGAGPMAADNNEGADGVNAGSNWCHDSITWLDGRIYITSTRTALPTYNHLY 120
DB 198 VGPPTWASGGAGPMAADNNEGADGVNAGSNWCHDSITWLDGRIYITSTRTALPTYNHLY 257
QY 121 KOISSASTGASNDNHFGYSTWPGYDFNRPHCHSPRPMORLINNNWGRPRRLNFKLF 180
DB 258 KOISSASTGASNDNHFGYSTWPGYDFNRPHCHSPRPMORLINNNWGRPRRLNFKLF 317
QY 181 NIOVEKVTNDGTTTANNLSTVOVFSSEYOLPVLGSAHOGCLPPPADVFMIPQYG 240
DB 318 NIOVEKVTNDGTTTANNLSTVOVFSSEYOLPVLGSAHOGCLPPPADVFMIPQYG 377
QY 241 YLTLLNGSOAVGRSSFYCLEYFPSPQMLRTGNFTFSYTEEVPFHSYAHOSGLDRLAMP 300
DB 378 YLTLLNGSOAVGRSSFYCLEYFPSPQMLRTGNFTFSYTEEVPFHSYAHOSGLDRLAMP 437
QY 301 LIDQVLYLNRITQNGSAGNKKDLFRSGSPAGMSVQPKMWLPGCYRQORVSKTKTDNN 360
DB 438 LIDQVLYLNRITQNGSAGNKKDLFRSGSPAGMSVQPKMWLPGCYRQORVSKTKTDNN 497
QY 361 NSNFTWTGASKYNLNGRESITNGTAMASHKDEDEFPFSGVMTFGKESAGSANTALDN 420
DB 498 NSNFTWTGASKYNLNGRESITNGTAMASHKDEDEFPFSGVMTFGKESAGSANTALDN 557
QY 421 WMTDEBEIKATNPVATERFGTAAVNFOSSTDPATGDVYAMGALPGMWQDRDYLQGP 480
DB 558 WMTDEBEIKATNPVATERFGTAAVNFOSSTDPATGDVYAMGALPGMWQDRDYLQGP 617
QY 481 WMTDEBEIKATNPVATERFGTAAVNFOSSTDPATGDVYAMGALPGMWQDRDYLQGP 540
DB 618 WMTDEBEIKATNPVATERFGTAAVNFOSSTDPATGDVYAMGALPGMWQDRDYLQGP 677
QY 541 QVSVEIEMELQKNSKRNAPVEQYTSNYAKSANVDFVNDNGLYTEPRPRTGTYLRLPL 599
DB 678 QVSVEIEMELQKNSKRNAPVEQYTSNYAKSANVDFVNDNGLYTEPRPRTGTYLRLPL 736

RESULT 15
ADV70293
ID ADV70293 standard; protein: 736 AA.
XX
AC ADV70293;
XX
DT 10-MAR-2005 (first entry)
XX
DE Primate adeno-associated virus 6 capsid protein VP1.
XX
KM immunosuppressive; gene therapy; immunogenicity; virus inactivation;
KM hemophilia; Peppock deficiency; galactosemia; phenylketonuria;
KM Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
KM atherosclerosis; thrombosis; embolism; Parkinson's disease;
KM congestive heart failure; cancer; inflammation; immune disorder;
KM muscular dystrophy; diabetes; VPI.
XX
OS Adeno-associated virus 6.
XX
PN WO2004112727-A2.
XX
PD 29-DEC-2004.
XX
PF 21-JUN-2004; 2004MO-US019884.
XX
PR 19-JUN-2003; 2003US-0480395P.
PR 30-APR-2004; 2004US-0567310P.
PR 03-JUN-2004; 2004US-0576501P.
XX
PA (AVIG-) AVIGEN INC.
PI Arbetman AE, Colosi P, Lochrie MA, Surosky RT;
XX
DR WPI; 2005-048755/05.

XX New mutated adeno-associated virus (AAV) capsid protein that when present
PT in an AAV virion imparts decreased immunoreactivity to the virion as
PT compared to the corresponding wild-type virion, useful for treating e.g.
PT hemophilia.
XX
PS Example 5; SEQ ID NO 19; 136pp; English.
XX
CC The invention describes a mutated adeno-associated virus (AAV) capsid
CC protein that when present in an AAV virion imparts decreased
CC immunoreactivity to the virion as compared to the corresponding wild-type
CC virion. Also described are: a polynucleotide encoding the mutated protein
CC above; a recombinant AAV virion comprising the mutated protein above; and
CC delivering a recombinant AAV virion to a cell or tissue of a vertebrate
CC subject. The recombinant AAV virion is useful for delivering a
CC heterologous nucleic acid molecule to a cell or tissue of a vertebrate
CC subject, where the protein encoded by the heterologous nucleic acid
CC molecule is expressed at a level that provides a therapeutic effect,
CC where the recombinant AAV virion may comprise a non-primate, mammalian
CC AAV capsid protein that when present in an AAV virion imparts decreased
CC immunoreactivity to the virion as compared to immunoreactivity of primate
CC AAV-2, and the heterologous nucleic acid molecule, where the heterologous
CC nucleic acid molecule encodes a therapeutic protein and is operably
CC linked to control elements capable of directing the in vivo transcription
CC and translation of the protein. The protein or the recombinant AAV virion
CC is useful for treating or preventing a wide variety of disorders such as
CC hemophilia, glycogen storage deficiency type Ia, Peppock deficiency,
CC galactosemia, phenylketonuria, Crigler-Najjar disease, Gout and Lesch-
CC Nyan syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,
CC thrombosis, embolism, Parkinson's disease, congestive heart failure,
CC cancer, inflammatory and immune disorders, muscular dystrophies, and
CC diabetes. This is the amino acid sequence of adeno-associated virus 6
CC (AAV6) capsid protein VP1.
XX
SQ Sequence 736 AA:
Query Match 99.3%; Score 3229; DB 9; Length 736;
Best Local Similarity 99.2%; Pred. No. 8.3e-252;
Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQDSESVDPQPLGEPATPPA 60
DB 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQDSESVDPQPLGEPATPPA 197
QY 61 VGPPTWASGGAGPMAADNNEGADGVNAGSNWCHDSITWLDGRIYITSTRTALPTYNHLY 120
DB 198 VGPPTWASGGAGPMAADNNEGADGVNAGSNWCHDSITWLDGRIYITSTRTALPTYNHLY 257
QY 121 KOISSASTGASNDNHFGYSTWPGYDFNRPHCHSPRPMORLINNNWGRPRRLNFKLF 180
DB 258 KOISSASTGASNDNHFGYSTWPGYDFNRPHCHSPRPMORLINNNWGRPRRLNFKLF 317
QY 181 NIOVEKVTNDGTTTANNLSTVOVFSSEYOLPVLGSAHOGCLPPPADVFMIPQYG 240
DB 318 NIOVEKVTNDGTTTANNLSTVOVFSSEYOLPVLGSAHOGCLPPPADVFMIPQYG 377
QY 241 YLTLLNGSOAVGRSSFYCLEYFPSPQMLRTGNFTFSYTEEVPFHSYAHOSGLDRLAMP 300
DB 378 YLTLLNGSOAVGRSSFYCLEYFPSPQMLRTGNFTFSYTEEVPFHSYAHOSGLDRLAMP 437
QY 301 LIDQVLYLNRITQNGSAGNKKDLFRSGSPAGMSVQPKMWLPGCYRQORVSKTKTDNN 360
DB 438 LIDQVLYLNRITQNGSAGNKKDLFRSGSPAGMSVQPKMWLPGCYRQORVSKTKTDNN 497
QY 361 NSNFTWTGASKYNLNGRESITNGTAMASHKDEDEFPFSGVMTFGKESAGSANTALDN 420
DB 498 NSNFTWTGASKYNLNGRESITNGTAMASHKDEDEFPFSGVMTFGKESAGSANTALDN 557
QY 421 WMTDEBEIKATNPVATERFGTAAVNFOSSTDPATGDVYAMGALPGMWQDRDYLQGP 480
DB 558 WMTDEBEIKATNPVATERFGTAAVNFOSSTDPATGDVYAMGALPGMWQDRDYLQGP 617
QY 481 WMTDEBEIKATNPVATERFGTAAVNFOSSTDPATGDVYAMGALPGMWQDRDYLQGP 540

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:29:19 ; Search time 21.1525 Seconds
(without alignment)
2724.684 Million cell updates/sec

Title: US-10-696-282-15

Sequence: 1 TAPGKKRPVEQSPQSPDSSS.....NNGLYTEPRPIGTRVLTPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_80:1*
1: pirl:1*
2: pirl:2*
3: pirl:3*
4: pirl:4*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2177.5	67.0	504	1 VCPV3A	coat protein - ade
2	1701	52.3	732	2 S52210	coat protein VP1 -
3	556	17.1	673	1 VCPVBS	coat protein VP1 -
4	495	15.2	781	1 VCPV19	coat protein VP1 -
5	258.5	8.0	723	1 VCPVPP	coat protein VP1 -
6	254.5	7.8	729	1 A60006	coat protein VP1 -
7	254.5	7.5	729	1 VCPVNA	coat protein VP1 -
8	244	7.2	722	1 VCPV2	coat protein VP1 -
9	234	7.2	587	1 B44276	coat protein VP1 -
10	216	6.6	727	1 VCPV2M	coat protein VP1 -
11	216	6.6	727	1 VCPV1F	coat protein VP1 -
12	214.5	6.6	718	1 VCPVIM	coat protein VP1 -
13	209	6.4	727	1 VCPVME	coat protein VP1 -
14	208	6.4	727	1 VCPVPP	coat protein VP1 -
15	208	6.4	737	1 VCPVND	coat protein VP1 -
16	204	6.3	748	1 VCPVCP	coat protein VP1 -
17	198	6.1	584	2 S49594	capsid protein VP2
18	192	5.9	722	1 VCPVNC	coat protein VP1 -
19	191	5.9	702	1 VCPVAP	coat protein VP1 -
20	143.5	4.4	648	2 S50856	whn protein - rat
21	129	4.0	1072	2 A86827	hypothetical prote
22	123.5	3.8	690	2 AB0124	probable Tomb-depe
23	123	3.8	1142	2 T37455	enamelin precursor
24	122.5	3.8	880	2 SYBSVS	valine-tRNA ligase
25	121	3.7	635	2 F96660	protein F2X1.10
26	121	3.7	1113	2 S28925	nuclear pore compl
27	120.5	3.7	931	2 T49710	related to gluan
28	119	3.7	667	2 A41311	transcription fact
29	118.5	3.6	1180	2 B86719	hypothetical prote

30	118	3.6	739	2 T52289	probable transekeo
31	116.5	3.6	642	1 S34416	transcription fact
32	116.5	3.6	1338	2 T30565	MAP kinase kinase
33	115.5	3.6	345	1 G97024	probable phosphos
34	114	3.5	956	2 T08144	myrosinase-binding
35	114	3.5	2271	2 F90073	hypothetical prote
36	113.5	3.5	2817	2 B97033	uncharacterized pr
37	113	3.5	1071	2 E85343	hypothetical prote
38	112.5	3.5	666	2 T52648	class A helix-loop
39	112.5	3.5	3078	2 T28432	variant-specific B
40	111	3.4	1777	2 T34369	hypothetical prote
41	110.5	3.4	655	1 ALKRG	cyclomalitodextrin
42	110.5	3.4	1742	2 T17120	cellulase (EC 3.2.
43	110	3.4	868	2 A82521	conserved hypothet
44	110	3.4	1296	2 C82521	hemolysin-type cal
45	110	3.4	2825	2 T14271	Doc4 protein, stre

ALIGNMENTS

RESULT 1

VCPV3A
coat protein - adeno-associated virus type 2

C/Species: adeno-associated virus type 2
C/Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004

C/Accession: A03698
R/Srivastava, A.; Lusby, B.W.; Berns, K.I.

J. Virol. 45, 555-564, 1983

A/Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.

A/Reference number: A03694; PMID:83164299; PMID:6300419

A/Accession: A03698

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-504 <SRI>

A/Suprafamily: adeno-associated virus coat protein

C/Keywords: coat protein

Query Match	Score	Length	DB 1	Length	DB 2
66	MASGGAPMADNNEGADGVMASGNHCDSTWLGDRVITTSRTWMLPTVNNHLYQISS	125			
1	MATGSGAPMADNNEGADGVMASGNHCDSTWLGDRVITTSRTWMLPTVNNHLYQISS	60			
126	ASTGASNDNHVREGSTPMGVPDPNRPCHSPSPDQORLNNNGFPRKRLNFKLFIQVK	185			
61	QS-GASNDNHVREGSTPMGVPDPNRPCHSPSPDQORLNNNGFPRKRLNFKLFIQVK	119			
186	EVTINDGVTTIANNLSTVQVPSDSEYQLEPVYLSAHOGLPPPADVPMIPOYGYLTIN	245			
120	EVTINDGVTTIANNLSTVQVPSDSEYQLEPVYLSAHOGLPPPADVPMIPOYGYLTIN	179			
246	NSQAVGRSSFYCLEFPPSQMLRTGNFTSTYFEEVPRHSSVAHSQSLDLRLNPLIDQY	305			
180	NSQAVGRSSFYCLEFPPSQMLRTGNFTSTYFEEVPRHSSVAHSQSLDLRLNPLIDQY	239			
306	LYTLNRTQNSGSAOKKDLFSRGSPPAGMSVQPKMLPGPCYQQRVSKTTDNNNSNFT	365			
240	LYTLNRTQNSGSAOKKDLFSRGSPPAGMSVQPKMLPGPCYQQRVSKTTDNNNSNFT	299			
366	WTGASKYNLNGRESINPGTAMASHKDDDEKPPMSGVMIFGKESAGASTALDNVITD	425			
300	WTGATKYNLNGRSLNVP--AMASHKDDDEKPPMSGVMIFGKESAGASTALDNVITD	357			
426	EEBIIKATNPVATRFGTAVANFQSSSTDPATGVHAMGALPGWVWDORVYLGQPIWAKI	465			
358	EEBIIKATNPVATRFGTAVANFQSSSTDPATGVHAMGALPGWVWDORVYLGQPIWAKI	417			
486	PHDGHHPSPPLMGGRGLKXPPQILIKTTPVAPNPAPESATKFSFITQYSGT	540			

DB 418 PHTDGHFHSPLMGFGGLKHPPQILIKNTPVANPSTTFSAKAFSTOYSTG 472

RESULT 2

SS2210
coat protein VP1 - muscovy duck parvovirus
N:Alternate names: VP1 protein
C:Species: muscovy duck parvovirus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: SS2210
R:Zadori, Z.; Erdel, J.; Nagy, J.; Kisari, J.
submitted to the EMBL Data Library, September 1993
A:Reference number: SS2209
A:Accession: SS2210
A:Molecule type: DNA
A:Residues: 1-732 <ZAD>
A:Cross-references: UNIPROT:O83289; UNIPARC:UPI000006FCSD; EMBL:X75093; NID:g609091; PTD
A:Experimental source: strain FM
C:Gene: VP1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 52.3%; Score 1701; DB 2; Length 732;
Best Local Similarity 53.0%; Pred. No. 1.6e-108;
Matches 325; Conservative 83; Mismatches 165; Indels 40; Gaps 11;

QY 6 KRPEQSPQEPDSSSGIGK-TGQQPAKKRLNFGQTGDSSEVPDPPPLG-----E 53
DB 141 EBPVNTAPAKKSS-----GKLTDDHPVKKPKLSE-----ENSPSPNSGGESAALATGSE 192
QY 54 PATPAVPTTMAAGGAPMADNNGADGVGNAGNHCDSWTGDRVITTSRTMALP 113
DB 193 PVAAP-----NMAAGSGAGMGDSAGAGDVGAGNAGNHCDSWTGDRVITTSRTMALP 246
QY 114 TNNHLVKYKISSASTGASNDNHYFGYSTPMGYFDFNRFCHFSRPMQRLINNNMGFRPK 173
DB 247 STNNHMYKAITSGTNPDSN-TOYAGYSTPMGYFDFNRFCHFSRPMQRLINNNMGIRPK 305
QY 174 RLNFKLFNIOVKVETTTNDGVTIANLSTVQVPSDSEVQLPYVLGSAHQGLPPPADV 233
DB 306 ALKFKIFNVQVKEVTTQDQTKTIANNLTSTIQIFTDNEHQLPYVLGSAHQGLPPPADV 365
QY 234 FNIPOGYVTLN---NGSAVGRSSPYCLEYFSPQMLRTGNNFTSSYFEERYEPHSSVAH 290
DB 366 YALPOGYCTMHTNOSGARFNDRSAYCLEYFSPQMLRTGNNFESFEERYEPHSMFAH 425
QY 291 SGLRLNMPILIDQYLYYLNRTONOSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQ 350
DB 426 SODLRLNMPILIDQYLMNFSV-NGGRNAQ-----FKKAVKGAFGAMGRNMLPGPKLDD 479
QY 351 RVSKTK--TDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIRGK 408
DB 480 RVRAVSGGTDNYANMSIGKGNKVFLEKDEYLLQPPVATHTTEQDASSVPAQNIIIGLAK 539
QY 409 E--SAGASNTALDNWITDEEELKATNPATRFQVAVNPOSSSTDPATGVAHMGALP 466
DB 540 DRYRSGSTLAGISDINVTDEQELAPTNVGMRPYGLVTNQNQNTTAPNAELBYVGLAP 599
QY 467 GMYMODRVYLOGPIWAKIPHTDGHFHSPLMGFGGLKKNPPOILIKNTPVANPPAPFS 526
DB 600 GMYMODRVYLOGPIWAKIPKTDGKPHSPNIGRGLNHPPOVFIKNTPVADPPLEAV 659
QY 527 ATKFASTFOYSTGQVSEIEMELQKNSKRANPEVOYTSNAYKASANDFTVDNNGLYTE 586
DB 660 NQKMSYITQYSTGCTVEMVWELRKENSKRANPEIQFTSNFGNRTSTWFAENETGYVE 719
QY 587 PRPIGRTYITRPL 599
DB 720 DRLIGTRLYTONL 732

RESULT 3

VCPVB5
coat protein VP1 - bovine parvovirus

N:Contains: coat protein VP2
C:Species: bovine parvovirus
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A26104
R:Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.
J Virol. 60, 1085-1097, 1986
A>Title: Complete nucleotide sequence and genome organization of bovine parvovirus.
A:Reference number: A26104; MUID:87061184; PMID:3785814
A:Accession: A26104
A:Molecule type: DNA
A:Residues: 1-673 <CHB>
A:Cross-references: UNIPROT:P07297; UNIPARC:UPI0000127D77; EMBL:M14363; NID:g333454; PTC
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:138-673/Product: coat protein VP2 #status predicted <VP2>

Query Match 17.1%; Score 556; DB 1; Length 673;
Best Local Similarity 25.0%; Pred. No. 3.4e-30;
Matches 160; Conservative 98; Mismatches 239; Indels 144; Gaps 19;

QY 1 TAPGKRPVQSPQEPDSSSGIGKTKGQPAKKRLNFGQTGDSSEVPDPPPLGEP--PATP 58
DB 100 TSKGGDRALKKRLYFARSNGAKKANREPASTNSQNMVEVNDIPNDEAGNQTIELATR 159
QY 59 AAVGPTTMAAGGAPMADNNGADGVGNAGNHCDSWTGDRVITTSRTMALPTYNNH 118
DB 160 SVAGSGVSGGGG-----RGSGGVGYSTGWTGQGTTFSENIVYTKNTROFICDINKGH 211
QY 119 LYKQISSASTGASNDNHYFGYSTPMGYFDFNRFCHFSRPMQRLINNNMGFRPKRLNFK 178
DB 212 LYKS-EVANTDPTAHQY-ALTPMSYFNFMQYSHFSPNDMQLVNDYERFRPAKMLVR 269
QY 179 LENIOVKVETTTNDGVTIANL-LTSTVQVPSDSEVQLPYVLGSAHQGLPPPADVEMIP 237
DB 270 VYNIQIKQIMTDGAMGTYYNNDLITAGMHI FCDGHRHYLYVQHPMDQCMPLPNSIWEIP 329
QY 228 QYGYITL-----NNGSAVGRS-----SFYCLRFSPQMLRTGNNFFSTFEEVPHS 286
DB 330 QYAYIPAPISVNDNTTNTVBEHLKGVPLYLENSDHEVLNNG----- 373
QY 287 SVASQSLRLNMPILIDQYLYYLNRTON-----QSGSAQNKDLLFSRGSF-- 331
DB 374 -----RKYRIYQMLRMDRKHQIHASDDVOSTGOKQNLIIQTKQPNKQRF 424
QY 332 AGMSVQPKNMLPGPCYRQVRSKTKTNDNNNSNFTWTGASKYNLNGRESIINPGTAMASHK 391
DB 425 ONAALRTSNMMSGP-----GIARGTNATLQTOGAGALVTWVT 462
QY 392 DDEDKFFPMGVM-----IFGKE-----SAGASNTALDNWITDE 426
DB 463 NGAD-----VSGVRAVRGYSTDPYIGGQPPSLLRLRLYSASAAEGQNPILN----- 512
QY 427 EIKATNPATRFQVAVNPOSSSTDPATGVAHMGALPGMYMODRVYLOGPIWAKIP 486
DB 513 -----AAHHTFREARTKLITGNSGADGDEKEMMLPLPQNMOSAPISRNPLIMVAVP 564
QY 487 HTDGHFHSPLMGFGGLKKNPPOILIK--NTPVPANPPAPFSATKFASTFOYSTGQVSV 544
DB 565 RVNRKRTLLDQDGSIPMSHPTOTIFIKLARIIPVPGND-----SFLNIYVYGVQVSC 615
QY 545 ELEMELQKNSKRANPEVOYTSNAYKASAND-FTVDNNGLY 584
DB 616 EVWMEVEKRGTKNMRPEYTHS---ATNMSVAYATYTNAGVY 653

RESULT 4

VCPV19
coat protein VP1 - parvovirus B19 (strain Au)
C:Species: parvovirus B19
A>Note: host Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004

C:Accession: A24299
R.Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A:Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr
A:Reference number: A24299; MUID:86200451; PMID:3701931
A:Accession: A24299
A:Molecule type: DNA
A:Residues: 1-781 <SHA>
A:Cross-references: UNIPROT:P07299; UNIPARC:UPI0000127D7C; EMBL:ML3178; NID:G333375; PID
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 15.2%; Score 495; DB 1; Length 781;
Best Local Similarity 27.3%; Pred. No. 6.2e-26;
Matches 148; Conservative 84; Mismatches 257; Indels 54; Gaps 16;

```
QY 51 LGEPRTPAVPTTMAAGCGAPMADNNGADVGAGSNMHCDSFTWLDRTTSTRW 110
DB 213 LREVPAYNSERYKPSMTSVNSAB-ASTGAGGGGSSNVKSMSEBGFSSANVTCTFSKOF 271
QY 111 ALPTYNHLVKOISSASTGASND-----NHYFGSTPMGYDFENRFCHSPRDM 160
DB 272 LIPYDEHHYKVPSPASSCHNASKGKAVCTISPMGYSTRVYLDFAALMLFESPLF 331
QY 161 QRLINNMGRPRKLNFLKFNIOVKEVT--TNDGVTIANNLSTVQVFSDBYQLPYVL 218
DB 332 QHLIENYGIADPDLITVITSEIAVKQVDTKGGV-QVTDSTTGRLCMLVDHEKYPVL 390
QY 219 GSAHOGCLPPPADVPMITQYGLTIN-GSQAVG-----RSFYCLEYFPGQMR 268
DB 391 GGGODTLABELPIWVFPPOYALITVGDVNTGIGSDSKLASSESAFYVLEHSSPDLG 450
QY 269 TGNNTFFSTFEVPFPHSSVAHSQSLDRLMNPLIDQYLYLNRTQSGSQAQNKDLIFER 328
DB 451 TGGTASMSKTFPPVPEENIEGCSQHFEYMNPL---YGSRLGVPTDLGGDPKRSI- 503
QY 329 GSPAGNSVQPKWMLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGESIIINPG- 387
DB 504 -THEDIAIQPNMPPGLVNSVSTKEGDSNTGAKALITGLSTGTSQNTRIISLRPGVQ 562
QY 388 ASHKDEDEKFFPMGSGVMIFGKESAGSNALDNV-----MTDEEIKATPVATERR 440
DB 563 PYHNMWTDXYVGINAISHGQTTYG--NAEDKEYQGVGRFNEKEOLKOLGMLHMTY 619
QY 441 GTAAVNFQSSSTDPATGDVYAMGALPGMWODRDVYLQGPIMAKIPIHTGCHFHPS- 499
DB 620 -----FPNKGITQYTDQIE-RPLMVGSVNRRALMYESQMSKINLSDSFITQFAALG 672
QY 500 GFGLKNPPQIILKNTVPVNPAPPAEFSAITKFSFITQSTGVSEIEMEL-OKENSKEM 558
DB 673 GMLGHPRPQIPLK--ILPQSGPIGGIKSMGITTLVQYANAGIMVTMTFGLGPKATGRN 730
QY 559 NPE 561
DB 731 NPQ 733
```

RESULT 5

coat protein VP1 - porcine parvovirus
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
R/Ranz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casael, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A:Title: Porcine parvovirus: DNA sequence and genome organization.
A:Reference number: A33302; MUID:90010964; PMID:2794971
A:Accession: B33302
A:Molecule type: DNA
A:Residues: 1-723 <RAN>
A:Cross-references: UNIPROT:P18546; UNIPARC:UPI0000174964; EMBL:D00623
C:Genetics:

A:Introns: 10/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match 8.0%; Score 258.5; DB 1; Length 723;
Best Local Similarity 23.1%; Pred. No. 8.6e-10;
Matches 155; Conservative 92; Mismatches 270; Indels 153; Gaps 35;

```
QY 9 VEQSPQEDSSSGIGK-----TGOQPKKRLNFGQGDSESVDPDPLGEPPTP 58
DB 103 VRSPRKHGKSPGPKRPAAPHIFINLAKKAKGTSNTNSMSNENVOHPIN--AAE 160
QY 59 -AAVPTTMAAGCGAPMADNNGADVGAGSNMHCDSFTWLDRTTSTRW 112
DB 161 LSATNBSGGGGG-----GGRGAGGVSTGSPFNQGTPEFQYLGGLVITMAHSLIHL 216
QY 113 PTYNHLVKO-----SSASTGAS-NDNHYFGSTPMGYDFENRFCHSPRDMOQLINN 166
DB 217 NMPEHETTKRLIHLNLSBSGSAQVQDDAHQMTVPMSLIDANMAGWFMNPADMQLISN 276
QY 167 NMGRPRKLNFLKFNIOVKEVT--TNDGVTIANNLSTVQVFSDBYQLPYVLGSAHQ 223
DB 277 MTEINLVSPQQAIFNVVLKTTESATSPPTKIYNNDLTASLWALDNTNLTLPYPAAPRS 336
QY 224 GCLPPPADVPMITQYGLTIN-GSQAVG-----TLNGSQAVGRSS-----FYCLE-YRP 263
DB 337 ETLGFPYPLPTKPTQYRYLLSCIRMLNPPTYGQSPNNRLNTRLHSDIMFYTIENAVP 396
QY 264 SGMRTGNPF--TFSTFEVPFPHSSVAHSQSLDR-----LAMPPLI--DOYLYLNRT 312
DB 397 IHLRTGDEFSTGIYHFDTKPL--KLTHSWQNRSLGLPRLKLTPTTGGDHPGTLPA 454
QY 313 QNQGSAQNKDLIFSRGSPAGNSVQPKWMLPGPCYRQQRVSKTKTDNNNSNFTWTGASKY 372
DB 455 NTRKYVHTQINNSYREAT--AIRP-----AQGVYTPMYNFEYSNGCF 496
QY 373 NNGESIIINP--GTMAASHKDEDEKFFPMGSGVMIFGKESAGSNALDNV-----NWMTTDE 427
DB 497 -----LTPVTPADTQVNDENP-----GAIRFTMYDGHLLTSSQ 533
QY 428 EIK--ATNP-----VATERPGTVA-VNPOSS-----TDPATG--DVHAGALP----- 466
DB 534 ELERYTFNPOSGKGAAPKQAFNOQAPLNLNNTNGLTLPBIDYIGKSNHFMNTLNTYGP 593
QY 467 -GMVWODRDVYLQGPIMAKIPIHTD--GHPHSPMLMGFGLK-NPPQIILKNTVPVNPAP 522
DB 594 LTAALNTAPVFPNGQIMDELDLTKPRH--VTAFFVCQNNPPQQLFVXIAP--NL 647
QY 523 AEFSA-TKFAFITQYSTGVSEIEMELOKENSKRANPEVOYTSYAKSANDPFTVDNN 581
DB 648 DDFNADSPQQRPIITYSNFMWKGTLTFTAKMRSSNMWNIQOHTT-----TAENI 697
QY 582 GLYTERPIG 591
DB 698 GXYI-PTNIG 706
```

RESULT 6

coat protein VP1 - porcine parvovirus (strain 90HS)
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
R/Sakurai, M.; Nishimori, T.; Ushimi, C.; Nakajima, H.
Virus Res. 13, 79-86, 1989
A:Title: Nucleotide sequence of capsid protein gene of porcine parvovirus.
A:Reference number: A60006; MUID:89319168; PMID:2750278
A:Accession: A60006
A:Molecule type: DNA
A:Residues: 1-729 <SAK>
A:Cross-references: UNIPROT:P33484; UNIPARC:UPI0000127C66

C:Superfamily: parvovirus coat protein
C:Keywords: coat protein; glycoprotein
E:151-729/Product: coat protein VP2 #status predicted <VP2>
F:172, 158, 282, 330, 433, 471, 573, 604, 551/Binding site: carbohydrate (Aan) (covalent) #status

Query Match	7.8%;	Score 254.5;	DB 1;	Length 729;
Best Local Similarity	22.7%;	Pred. No. 1.6e-09;		
Matches 151;	Conservative 95;	Mismatches 277;	Indels 141;	Gaps 33;

```

0Y 9 EBOSQOEEDSSGDIK-----TGQPAKKRLINFGQTDGSEBVPQPLGEBPAP 58
Db 109 VRSRPKRHPGSRPPGRRPAPRPHI F1NLAKKKAGKSTNNSKSENBEOHNPINACTEL- 167
0Y 59 AAVGPTTASGGGAPPADNNEGADVGNASGMWH--CDSTWLGD---RYITSTRTWPAL 113
Db 168 SATNGESGGGGGG-----GGRGAGGVGTSGSFNNQTEFOYLGBGLVRIATAHSRLIHLN 223
0Y 114 TYNNHLKYOS--SASTGAS---NDNHFGYSTPMGCTFDNRFHCHFSBDMQRLINN 167
Db 224 MPEHTYKRIHVLNBSGVAQGMVODDATQMTWTPSLIDAAMGWFENPADQLLSNM 283
0Y 168 WGFREKRLNFKLFENICVKEVT---TWQDTTIANNTSTYQVSDSEYOLPYLGAHQG 224
Db 284 TEINLVSEPOELFNVVLKTTITESAISPPKTYNNDLTALMALDINNLTLPYPAPRBE 343
0Y 225 CLPPPADVFMIPQGYL-----TLNNGQAVGRS-----SFYCLB-YFPG 264
Db 344 TLGFPWMLPTKPTQYXYULSTRNILNPPYTGQSOQITDSIQGLHSDIMFYIENAVPI 403
0Y 265 OMLRTGNNP--TFSYPEEVPFHSVYASQSLD-----LMPPLI--DOYLYULNRTQ 313
Db 404 HILRTGDESTGTYHFDTKPL--KLTHSQTRSLDPRKLTLPTEBGDQHCTPAPAN 461
0Y 314 NQSGAQNKDLLFSRGSPPAGMSVOPRNMLPGCYROORVSKTKTDNNNSFTWTGASKYN 373
Db 462 TRKGHYQTNNNSYTEAT--AIRP-----AQVGYNTPYNNFEYSNGGPF--502
0Y 374 LNGRESSINP--GTMAASHKDEDEKFPMSGVMI POKESAGSNTALDVMITDEBEICA 431
Db 503 -----LPIVPTADTOYNDDE---PNCALFTMGYOHGOLTS-----SOELERYT 545
0Y 432 TNP-----VATERFQTV--VNFQSGS-----TDPATG--DVHANGALP-----GMVWQ 471
Db 546 FVPQSKCGRAPQGFNQARPLNLENNNGTLLPSDITGGKPNHFNNTLNTYGRPLALNN 605
0Y 472 DRDVTYLOGPIAKIPIHTD--GHFHSPLMGGSFLK--NPPQILLIKTVPYAPNPAEFA- 527
Db 606 TAPVPPNGIOWKEJELDTLCKPRIH---VAPFCKNNPPOULPVKIALP---NLTDFEND 659
0Y 528 TKFASFITQYSTQGVAVELEMELOKENSKRMPNEVOYTSYVAKASAVUDFTVANGYTER 587
Db 660 SPQQRIRIITYSNFMWKGTLFTFAKMSNMNMPIOQHTT-----TAENIGYI-P 708
0Y 588 RPIG 591
Db 709 TWIG 712

```

```

RESULT 7
VCPVNA
coat protein VP1 - porcine parvovirus (strain NMDL-2)
N.Contains: coat protein VP2
C.Species: porcine parvovirus
C.Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C.Accession: B33743; D48472
R.Vaandevacharya, U.; Baask, S.; Striniwas, R.V.; Compans, R.W.
Virology 173, 368-377, 1989
A.Title: Nucleotide sequence analysis of the capsid genes and the right-hand terminal pa
A.Reference number: A33743; PMID:90085785; PMID:2596019
A.Accession: B33743
A.Molecule type: DNA
A.Residues: 1-729 <VAS>
A.Cross-references: UNIPROT:P18546; UNIPARC:UPI0000127670; GB:M33787; NID:G332983; PIDN:

```

R. Bergeron, J.; Menezes, J.; Tijssen, P. *Virology* 197, 86-98, 1993
 A1: title: Genomic organization and mapping of transcription and translation products of the
 A1: Reference number: A48472; MUID:94025614; PMID:8212598

A/Cross-references: UNIPARC:UPI0000174965
A/Cross-references: NADL-2, ATCC VR-742
A/Note: sequence extracted from NCBI backbone (NCBIN:138789, NCBIP:138794)

A: Introns: 10/1
C: Superfamily: parvovirus coat protein
E: Key words: coat protein; glycoprotein
F: 151.729/Product: coat protein; VP2 #status predicted <VP2>
F: 152.172/1.198,282,330,453,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #st

```

QY 9 ZEQSQOEDESSGCTK-----TQOQPAKKXKLNFEQGTDSSEVPPOPLGSEPARP 58
Db 109 VRBSRKHKPGSPKPKRPAHPRIFTINLAKKAKGTSNTNSMSSEVEQHNFINACTEL- 167
QY 59 AAVGPTTMASSGGGAGMADNNEGADVGNASG--NMHCDSITWLGD--RVYITSTRTALP 113
Db 168 SATGNESGGGGGGG---GGRGAGGVGVSTGTFFNNQTEFOYLGEGILVITTAHARSILHN 223
QY 114 TYNNHLXYQIS--SASTGAS---NDNHFGYSTMGVDFRFRHCHFBPDMOBLINN 167
Db 224 MEHEITYRIHVLNSESQVAGOMVODDATQMTVPISGLIDAAMGWEPFADQLISNM 283
QY 168 WGFREKRLNFKLENIQVKEVT---TNDVTTIANNLTSIVQYFSDBEYOLPYVLGSAHQ 224
Db 284 TEINVSFEQELFNNVLKTTESATSPPKIYNNDLTASLVALDNTNLTLPYPAARSE 343
QY 225 CLPPPADVFMIPQYGL-----TLNNGSQAVRS-----SFYCLE-YFBS 264
Db 344 TLGFPMWLPETKPYORYYLSCIRLNPPYYTQSOQITDSIOTGLHSDIMFYITENAVPI 403
QY 265 OMLRTGNPF-FPSYFEEVPFSSVAHSOSLDR-----LMPPLI--DOYLYUNRTO 313
Db 404 HILRTGDEFSIOIYHFDTKPL--KLTHSQMTRSLGLPKLLEPTTBDOHQHTLPAAN 461
QY 314 NQSGAQNKDILFSRSGSPAGMSVQPRNWLPGCFRQORVSKTKTDNNNSNFTWTGASKYN 373
Db 462 TRKGHQITINNSTYAT--AIRP-----AQGYNTPYNNFEFSNCGPR- 502
QY 374 LNGRSLINP--GTMASHKDEDEKFPFMSGVMIRGKESAGASNTALD--NMVIDEE 428
Db 503 -----LPIPIVLTDOYINDEPN-----GALRFMTDOHGLTSSDE 540
QY 429 IK-ATNP-----VATERFGTVA--VNFOSSS-----TDPATG--DVHAMGALP- 466
Db 541 LERYTFNPOSCKGARPKQOFNOQAPLNLNTNTNGLLPBDDPIGCKSMHFMNLTNYGPL 600
QY 467 GMYMODRDVYLOGPIYAKIPTH--GHFRPSLMGFGLK--MPPDILLKNTPVPANPPA 523
Db 601 TALNNTAPVPFNGQIMDELDLTKPRLL--VTAPEVCNNPFGOLFVKIAB--NLTD 654
QY 524 EESA-TKFASFITYQSTQGVSAVEIEMELOKENSKRNNVEQVTSNKAASANDFTVDNG 582
Db 655 DFNADSPOQPRITITSNFMWKGTLFTAGRSSNMNPIQOHTT-----TABNIG 704
QY 583 LYTERPPIG 591
Db 705 NYI-PTNIG 712

```

RESULT 8
VCPV2
coat protein VP1 - parvovirus H1

C.Species: parvovirus H1
A>Note: host Homo sapiens (man)
C.Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C.Accession: A03699
R:Rhode III, S.L.; Parodi, P.R.
J.Virol. 45, 173-184, 1983
A>Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybridization
A:Reference number: A03695; MUID:83112183; PMID:6823009
A:Accession: A03699
A:Molecule type: DNA
A:Residues: 1-722 <RHO>
A:Cross-references: UNIPROT:P03136; UNIPARC:UPI000012707D; EMBL:X01457; EMBL:J02198
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 7.5%; Score 244; DB 1; Length 722;
Best Local Similarity 21.4%; Pred. No. 8.4e-09;
Matches 142; Conservative 94; Mismatches 267; Indels 160; Gaps 29;

```
QY      8 PVEQSPQEPDSSSGIGKTGQO---PA-----KKRLNFGQTGDSSEVPDPPLGEP 54
      82 PKLSTDEP-GTSGVSRPGKRTKPAHIFVNGARAKKASLAQQRLLTMSDGETNP 140
QY      55 PATPAVGTMTASGGGAPMADNNEGADVGNAAGWCHDSTM--LGDVITTSRTMAL 112
      141 DTGIANARVERSDGGGS-SGGGSGGGIGVSTGYTDNQTYYKFGDGMVEITAHASRL 199
QY      113 -----PTYNHLYKQISSASTGASNDNHVFGYSTPMGYPDFNPFCHFSR 158
      200 LHLGMPSENYCRVYVHANNQTTGHGKVKGNMAVDTHQDIW-TPMSLVANAMGWVFOQS 258
QY      159 DMQRLINNMGPRPKLNFKLFNIQVKEYT-----NDGVTTIANLSTVOVFSDEYO 213
      259 DMQFIQNSMESLNDLSQELFVNVVKTTEQAGQADAIKYVNNDLTACMVALDSNNI 318
QY      214 LPLVLSAAGCL-----PPRPAD---VFMIPOGYLLNNGSQ-----ANG-----R 253
      319 LPLTPAQOTSETLGFYPMKPTAPAPRYFFEMPRQLSVTSSNSAEGTQITDTTIGEPOLN 378
QY      254 SSFYCLE-YFPSQMLRTGNFTF-SYTFEEVPHSSVA-----HSGSLDLNAN-PLIDQY 305
      379 SQFTLENTLPTLLTGTDEFTTGTYIFNTDPLKLHTWQTRHNLACLOGITDLPFSDBA 438
QY      306 LLYLNTQNGSGAONKDLF-----SRGSPAGNSVOPKMWL 342
      439 TABLTANGDRFGSTQONVNVYTEALRTRPAQIGFMQPHDNFEANGGPRKVVVP---- 494
QY      343 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGR--ESTINPCTA-----MASH 390
      495 -----LDITAGEHDHNDAN---GAIRENYGKHGEHWAKOGAAPERYTMDAIDSAG 541
QY      391 KODEDKPFPMGSGWIGKESAGASNTALDNVMTIDESEIKA-TNPVATERFGTVAVNFOS 449
      542 RDTARACV-----QSAPISIPNQNQILQREBALAGRTNMHTVNFNSYGPLSAR 591
QY      450 SSTDPATGDVHAMGALPGMWODRDVYLQGPIMAK--IPHTDGHFHPSPLMGFGGLKXP 506
      592 PHDP-----LYPNQIWDKELDEHKPRLLHVTAFV---CKNRP 628
QY      507 PPQILIKNTFVPANPAEF--SATKRFSTIOYSTGOVSVEIEMELOKENSKRNPEVOY 564
      629 PGOLFVHLGP---NLTDQFDPNSTVSRIVT--YSTFYWKGIKFKAKLRNLMTNPLYQA 684
QY      565 TSN 567
      685 TTD 687
```

RESULT 9
B44276
coat protein VP1 - parvovirus LuIII
C:Species: parvovirus LuIII
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004

C:Accession: B44276
R:Difford, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A>Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unit
A:Reference number: B44276; MUID:93297126; PMID:8517025
A:Accession: B44276
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-587 <DIR>
A:Cross-references: UNIPROT:P36310; UNIPARC:UPI000012707E; GB:M81888
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein; glycoprotein
F:49,90,220,304,371,503,511,514,539,571/Binding site: carbohydrate (Asn) (covalent) #act

Query Match 7.2%; Score 234; DB 1; Length 587;
Best Local Similarity 23.1%; Pred. No. 3e-08;
Matches 137; Conservative 74; Mismatches 231; Indels 152; Gaps 29;

```
QY      59 AAVGPTMASGGAGMADNNEGADVGNAAGWCHDS--TWLGD---RVITTSRTMAL- 112
      23 AADGPG--GSGGG-----GSGGGVGVSTGSDYNDQTHYKFLGDMVELTAYSTMVHLN 74
QY      113 -PTYNN-----HLKQISSASTGASNDNHVFGYSTPMGYPDFNPFCHFSRDMQRLIN 165
      75 MPKSENYCRVYVHANNQTTGHGKVKGNMAVDTHQDIW-TPMSLVANAMGWVFOQSDQYISN 133
QY      166 NMGPRPKLNFKLFNIQVKEYT-TNDGVTTIA--NLTSVVOVFSDEYOPLVLSA 221
      134 NMHINLSLDELNNVVIKVTTEQNTGAEALKVYNNDLTAAWVALDSNNILPTPALD 193
QY      222 HGGCL-----PPRPA-----DVMIPOGYLLNNGSOAVGRSSFYCLEYEP 263
      194 NQETLGFYPMKPTLSPYRYFSCRNLSVYKDEGITTDWGLASGLNSQFTIENTQ 253
QY      264 S-QMLRTGNF-TFSYTEEVPFSSYHSSOSLDRMLPLDQ-----YL 306
      254 RINLRTGDEVATGYVFPTEPDIRLHTWQTRHNLGQPOPIELPSSDPANATLFGARGYR 313
QY      307 VYLNRTN-----QSGAONKD-LFSPGSPAGNSVOPKMWLPG----- 344
      314 SGLTQIQGRNDVTEATRVRAQVGCQPHDNFETRAGFKVVPVADITQGLDHDANGS 373
QY      345 --PCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESINPTAMASHHDEDEKFPFMSG 402
      374 LRYTYDKOHGGSWASQONNDRYTW-DAVNYD-SGR----- 406
QY      403 VMIFGESAGASNTALDNVMTIDESEIKATNPATER--FGTVAVNFOS--SSTDPAFG 457
      407 -----WTNNCFQSVFPFSERP---ANQILTRDLAKTKTDIHFNAFNSYGPLRA 454
QY      458 DVHAMGALPGMWODRDVYLQGPIMAK--IPHTDGHFHPSPLMGFGGLKXNPPQILIKX 514
      455 FPH-----PAPIYPOQIWDKELDEHKPRLLHQAQFV---CKNNAFGQLLVRL 500
QY      515 TP--VPANPAEFSATKRFASFTIOYSTGOVSVEIEMELOKENSKRNPEVOYTS 566
      501 APMLTDQYDPSNNSISRIYVGFPEWKGLTLKAK--MRPNA-TNPVFOISA 550
```

RESULT 10
VCPV2M
coat protein VP1 - minute virus of mice
C:Species: minute virus of mice, murine parvovirus
C.Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A03700
R:Atwell, C.R.; Thomson, M.; Merchinsky, M.; Ward, D.C.
Nucleic Acids Res. 11, 999-1018, 1983
A>Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.
A:Reference number: A03696; MUID:83143341; PMID:6298737
A:Accession: A03700
A:Molecule type: DNA
A:Residues: 1-716 <AST>
A:Cross-references: UNIPROT:P03137; UNIPARC:UPI000012706D; EMBL:V01115

C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 6.6%; Score 216; DB 1; Length 716;
Best Local Similarity 20.8%; Pred. No. 6.9e-07;
Matches 148; Conservative 93; Mismatches 271; Indels 198; Gaps 34;

```

QY 8 PVEGSDPQEPSSSG:GKGTG-----QQPAKKRL-----NFGQT-GDSRSVDP 48
DB 82 PLATDSEB-GTSGVSRACKRTTPRAYIFINQARAKKLTSSAAQSSGTSBDSGPS 140
QY 49 QPLGEPRA--TPAVGPTTMAAGGAPMADNNEGADVGNASGNWCHDS--TWLGD--R 101
DB 141 GNAVSAARVERAADQPG--GSGGG-----GSGGGGCVSTGSTDQNGHYRFLGDGWE 192
QY 102 VTTTSTRTALPTYNHLYKQI-----SSASTASDNHIFGYSTWGYDFRFRFCH 154
DB 193 ITALATRLVHLMPKSENYCRIRVHNTTDTSVKGNMADDAHEQIWTPLSLVDANAMGV 252
QY 155 PGRDMQRLINNMGPRPKRLNPKLFNIQVKEVTND---GVTTIANLSTVQVFSDS 210
DB 253 LQPSDQYICNTMSQNLVSLDQEIFNVVLKTVTEODLGGQAIKIYNNDLTCMMVAVDS 312
QY 211 EYOLPYVLSAHOGLPPPADVFMIPQYGY-----LTNNGSQAV-----G 252
DB 313 NNILPYTPAANGMETLGFYPMKPTIASPYRYFCVDRDLSTYENQEGTVHNWGTGKG 372
QY 253 RSGFYCLEYFPG-QMLRTGNF--TFSYTFEEVPHSSVHQSGLDRLMN--PLIDQYLYX 308
DB 373 IPQFFIENTOOITLLRTGDEBFGATGYVFDTSV--KLTHWTQTRKOLQAPLSTF-- 427
QY 309 LARTQNGSQANKDLLFSGSPAGMSVQPKML-----PG 344
DB 428 -EADTDAGT-----LTQSSRHGTTQMGVNMVSEALITRAQVGFQCPHNDFAASRAG 480
QY 345 PCYQOQRYSKTIDNNNSFTWTGASKYMLNGRESIINPG-----TAMASHKDE 394
DB 481 P-FAAPKVPADITQGVDEKANGSVRSYQKHGEMASHGAPABERYTMDETSFGSGRDTK 539
QY 395 DKPFG-----PMSGVMIFGKESAGASNTALDNVMTDEBEIKATNPVATERGTYAV 445
DB 540 DGFISGAPLVVPPPLNGI-----LTNANPIGTKN-----DI 570
QY 446 NFGS--SSTDPAQGVHAMGALPGMWODRDVYLQSPIMAX---LPHTDGHFHPRLMG 500
DB 571 HSNVFNISGPIATASH-----PSVYIPQSQIMDKSLDLBKRLHITAPV-- 617
QY 501 PGLKPPPOILIKTP--VPANPPAFSATKPFITQYSTGVSEIEMELOKENSRRM 558
DB 618 -CKNNAFGQMLVRLGPNLIDQYDPNGATLSRIVTGTGTFPMKGLTMRALRA-----NTTW 672
QY 559 NBEVQYTSYAKSANDFTVDNNGLY---TEPRPIGT-----RYLTRPL 599
DB 673 NBYQVVSAR-----DNGNSYMSYTKMLPFTATGMQSGVPLITREV 711

```

RESULT 11

VCPVIF

coat protein VP1 - feline panleukopenia virus

N:Contains: coat protein VP2

C:Species: feline panleukopenia virus, FPLV

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004

C:Accession: A03701

R:Carlsbom, U.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.

J. Virol. 55, 574-587, 1985

A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parv

A:Reference number: A03697; MUID:85265017; PMID:2991581

A:Accession: A03701

A:Molecule type: DNA

A:Residues: 1-727 <CAR>

A:Cross-references: UNIPROT:P04864; UNIPARC:UPI0000127D55; EMBL:M10824; NID:G333474; PID

C:Genetics:

A:Introns: 11/1

C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

F:144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.6%; Score 216; DB 1; Length 727;
Best Local Similarity 20.2%; Pred. No. 7e-07;
Matches 139; Conservative 101; Mismatches 259; Indels 188; Gaps 34;

```

QY 7 RPEGSDPQEPSSSG:GKGTGQPAKKRLNFGQTDSSVP-----DQPLGEPRA--TPAA 60
DB 110 KPTRSKPPPHIFINLAK-----KKAAGAQVKRDNQAPMSDQAVQDQGPVNRERA 163
QY 61 VGPPTMASGGAPMADNNEGADVGNASGNWCHDS-----CDSTWLGDRVITSTRTALP 113
DB 164 TGSNGSGGGG-----CGSGGCVGISTGTENNGTEPFKLENGWV--EITANSSRLVHLN 215
QY 114 TYNHLYKQI-----SSASTASDNHIFGYSTWGYDFRFRFCHFSRDRQRLNN 166
DB 216 MPSENYKRVVNMNDTAVKGNMADDTVHQTVPMSLVANAMGVAFNPGDQQLVNT 275
QY 167 NMGRPKRLNPKLFNIQVKEVT---TNDGVTTIANLSTVQVPSDSEYOLPYVLSAHQ 223
DB 276 MSELHVSFEQELFNVLKTVSBSATQPPKTVNNDLTASLMVALDSNTMPTPAARSR 335
QY 224 GCL-----PPFPA-----DVFMIPQYGLTLNNGSQAVGRS-----SPFYCL 259
DB 336 ETLGFYPMKPTIIPPMKRYFPQMDRLTLPSh-----TGTSTPTNIYHGTDPDDQVFTI 389
QY 260 E-YPPSQMLRTGNF--TFSYTFEEVPHSSVHQSGLDRLMN--PLIDQYLYNRTONO 315
DB 390 ENSVPHLLRTGDEBFGATGFEFFDCKP--CRLLHWTQTRKALGLRP-----FLNSLPQS 440
QY 316 SGSAQNDDLFSRSPAGMS-----VOPKN-MLPGCYQQRYSK----- 354
DB 441 EGATNFGDIGVQOQRRKRVTOGNTDYITEATTIRPEAVGSAPIYSFEASTQGFKPI 500
QY 355 -----TKTDNNNSFTWTGASKY--NLNGRESIINPG-----AMASHKDEDKFPMS 401
DB 501 AAGRGAGQTDENQA---DQPRYAFGRHQKQKTTTGTEPERFTYIAHQGT----- 549
QY 402 GVMIFGKESAGASNTALD-NVMTDEBEIKATNPVATERGTYAVNFGSSSTDPAQDVH 460
DB 550 -----GRYPADWTONIFNLPTNDVNLPTDIDIG---GKTGINV--TWI FNTYGLT 598
QY 461 AMGALPGMWODRDVYLQSPIMAKIPTHGPHHSPLMGSGFLK-----NRP 507
DB 599 ALNNVP-----PYENGQIMDKXEPDT-----LKRLLHVAAPFYCONNCP 638
QY 508 PQLIKNTVPAN---PPAFSATKPFITQYSTGVSEIEMELOKENSRRNPEVOY 564
DB 639 GQLFVKVAPNLINEDPDAANMR-----LYTGDFFMWKGLVFKAKLRASHHTNPPIQM 694
QY 565 TSNYAKSANDFTVDNNGLYTEPRPIG 591
DB 695 SIN-----VDNQFNVL-PNNIG 710

```

RESULT 12

VCPVIM

coat protein VP1 - minute virus of mice (strain MMV1)

C:Species: minute virus of mice, murine parvovirus

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004

C:Accession: B23008; B29510

R:Sanli, R.; Mcmaster, G.K.; Hirt, B.

Nucleic Acids Res. 13, 3617-3633, 1985

A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous

A:Reference number: A23008; MUID:85242059; PMID:3855242

A:Accession: B23008

A:Molecule type: DNA

A:Residues: 1-718 <SAH>

A:Cross-references: UNIPROT:P07302; UNIPARC:UPI0000174963; EMBL:X02481; NID:g60918; PIDN

R:Sanli, C.R.; Gardiner, E.M.; Tattersall, P.

J. Virol. 57, 656-669, 1986

A:Molecule type: DNA
A:Residues: 1-727 <MAR>
A:Cross-references: UNIPROT:P24640; UNIPARC:UP10000127D56; GB:X55115
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.4%; Score 209; DB 1; Length 727;
Best Local Similarity 20.1%; Pred. No. 2,1e-06;
Matches 138; Conservative 101; Mismatches 260; Indels 188; Gaps 34;

```
QY 7 RPVEQSPQEPDSSGIGKTGQPAKKRLNFGQTDSSVP---DQPLGEPPA--TPAA 60
DB 110 KPTKRSKPPHIFINLAK-----KKKAGAGQVRDNLAPMSDGAVPDGGQPAVRNERRA 163
QY 61 VGPPTMASGGGAPMADNNEGADGVNAGSNMH-----CDSTLGDGVITTSRTALP 113
DB 164 TGSNGSGGGG-----GGSGGVGISTGTNNQTEKFELENGV--ETANSSRLVHLN 215
QY 114 TYNHLYKQI-----SSASTGASNDNHFGYSTPMGYFDFNRFHCHFSPRDMQRLINN 166
DB 216 MPESENTRRVYVNNMDKTAIVGNMALLDIIHQIVTPMSLVDANAMGWAFNPGDQQLVNT 275
QY 167 NMGFRPKRLNFKLENIQVKEVT--TNDGVTTIANNLTSVQVPSDSEYQLPYVLSAHQ 223
DB 276 MSELHLVSFEQEIFNVVLKTVSBSATQPTKYVNNDLTASLWALDNNMTMFTPAAMRS 335
QY 224 GCL-----PPPPA-----DVFMIPOGYVLTLLNNGSQAVGNS-----SPCYL 259
DB 336 ETLGTFYPMKPTIPTPMRWYFQMDRTLIPSH-----TGTSGTPTNIYHGTDPPDVQFTTI 389
QY 260 E-YPPSQMLRTGNF--TFSYTFEEVPFHSVAHSQSLDRLMN--PLIDQYLYLNRTONQ 315
DB 390 ENSVPLHLRTGDFATGTFPFDCRP--CRLLHWQTRALGLRP-----FLNSLPQS 440
QY 316 SGSAQNKDLLFSRGSPPAGMS-----VQPKN-WLPGCYRQORVSK-----354
DB 441 EGATPFGDIGVOQDRKRGVTOGMNTDYITBATIMRPAEYGSAPYSSFEASTQGFPTPI 500
QY 355 -----TKTDNNNSNFTMGASKY---NUNGRESIINPOT---AMASHKDEDEDFPMS 401
DB 501 AARGGAQTDENQA--DGDPRVAFGRHQGKTTTGTETPERFYIAHDT-----549
QY 402 GVMIFGKESAGASNTALD-NVMITDEBEIKATNPATERFGTVAVNFOSSSTDPATGVH 460
DB 550 -----GRYEGDMIONINFLPVINDNVLPLTDPIG-----GKTGINY--TNIFNYGPILT 598
QY 461 AMGALPGMYWQDRDYLLQGPITWAKIPIHTDGHFHPSPMLGGFGLK-----NPP 507
DB 599 ALNNVP-----PYYPNGQIWDKEFDTD-----LKPRLHVNAPFVQNNCP 638
QY 508 PQILIKNTKTPVPA--PPAFSATKFAFIIOYSTGOVSVEIEMWLOKENSRRNDEVOY 564
DB 639 GQLFVKVAPNLITNEYDPPASAMSR---ITVYSDPMWKGKLVFPAKLRASHHTMPLQOM 694
QY 565 TSNVAKSANVDFTVNNGLYTEPRPIG 591
DB 695 SIN-----VDNQFNYYV-PNNIG 710
```

RESULT 15
VCPVCD
coat protein VP1 - canine parvovirus (strain CPV-D)

N:Contains: coat protein VP2
C:Species: canine parvovirus, CPV
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #ext_change 09-Jul-2004
C:Accession: A31163
R:Parish, C.R.; Aguadro, C.F.; Carmichael, L.E.
Virology 166, 293-307, 1988
A:Title: Canine host range and a specific epitope map along with variant sequences in the
A:Reference number: A31163; MID:89020796; PMID:3176341
A:Molecule type: DNA

A:Residues: 1-737 <PAR>
A:Cross-references: UNIPROT:P17455; UNIPARC:UP10000127D7A; EMBL:M23255; NID:G333467; PID
C:Genetics:
A:Introns: 26/3
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:1584-737/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.4%; Score 208; DB 1; Length 737;
Best Local Similarity 19.8%; Pred. No. 2,5e-06;
Matches 137; Conservative 96; Mismatches 260; Indels 200; Gaps 31;

```
QY 7 RPVEQSPQEPDSSGIGKTGQPAKKRLNFGQTDSSVP---DQPLGEPPA--TPAA 60
DB 120 KPTKRSKPPHIFINLAK-----KKKAGAGQVRDNLAPMSDGAVPDGGQPAVRNERRA 173
QY 61 VGPPTMASGGGAPMADNNEGADGVNAGSNMH-----CDSTLGDGVITTSRTALP 113
DB 174 TGSNGSGGGG-----GGSGGVGISTGTNNQTEKFELENGV--ETANSSRLVHLN 225
QY 114 TYNHLYKQI-----SSASTGASNDNHFGYSTPMGYFDFNRFHCHFSPRDMQRLINN 166
DB 226 MPESENTRRVYVNNMDKTAIVGNMALLDIIHQIVTPMSLVDANAMGWAFNPGDQQLVNT 285
QY 167 NMGFRPKRLNFKLENIQVKEVT--TNDGVTTIANNLTSVQVPSDSEYQLPYVLSAHQ 223
DB 286 MSELHLVSFEQEIFNVVLKTVSBSATQPTKYVNNDLTASLWALDNNMTMFTPAAMRS 345
QY 224 GCL-----PPPPA-----DVFMIPOGYVLTLLNNGSQAVGNS-----SPCYL 259
DB 346 ETLGTFYPMKPTIPTPMRWYFQMDRTLIPSH-----TGTSGTPTNIYHGTDPPDVQFTTI 399
QY 260 E-YPPSQMLRTGNF--TFSYTFEEVPFHSVAHSQSLDRLMN--PLIDQYLYLNRTONQ 315
DB 400 ENSVPLHLRTGDFATGTFPFDCRP--CRLLHWQTRALGLRP-----FLNSLPQS 450
QY 316 SGSAQNKDLLFSRGSPPAGMSVQPKNWLPGCYRQORVSKTKTDNNNSNF-----TW 366
DB 451 EGATNFGDI-----GV-----QDDRKRGVTOGMNTDYITBATIMRPAE 488
QY 367 TGASKNLNGRESIINP-----GTAMASHKDEDEKFFMSVMTFGKESAGASNTALD-- 419
DB 489 VGYSAPYSSFEASTQGFPTPIAARGGAQTDENQAADGNRVAFGHQHGTCTTTGTP 548
QY 420 -----NVMITDEBEIKATNPATERFGTVAVNFOSSSTDP 454
DB 549 ERFTYIAHQDGRYEGDMIONINFLPVINDNVLPLTDPIG-----GKTGINY--TNIFN 602
QY 455 ATGDVHAMGALPGMYWQDRDYLLQGPITWAKIPIHTDGHFHPSPMLGGFGLK-----504
DB 603 TYGPILTALNNVP-----PYYPNGQIWDKEFDTD-----LKPRLHVNAPFV 642
QY 505 ---NPPQILIKNTKTPVPA--PPAFSATKFAFIIOYSTGOVSVEIEMWLOKENSRRN 558
DB 643 CONNCPQGLFVKVAPNLITNEYDPPASAMSR---ITVYSDPMWKGKLVFPAKLRASHHTM 698
QY 559 NPEVQTSNVAKSANVDFTVNNGLYTEPRPIG 591
DB 699 NPIQOMGIN-----VDNQFNYYV-PSNIG 720
```

Search completed: November 23, 2005, 17:44:02
Job time : 23.1525 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:25:34 ; Search time 115.377 Seconds
(without alignments)
3662.868 Million cell updates/sec

Title: US-10-696-282-15
Score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLRPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3251	100.0	736	2 Q9WB88_V9IRU	Q9WB88 adeno-ssoc
2	3230	99.4	736	2 Q6JC08_V9IRU	Q6JC08 adeno-ssoc
3	3230	99.4	736	2 Q6JC12_V9IRU	Q6JC12 adeno-ssoc
4	3229	99.3	736	2 Q6JC17_V9IRU	Q6JC17 adeno-ssoc
5	3203	98.5	736	2 Q6JC10_V9IRU	Q6JC10 adeno-ssoc
6	3163.5	97.3	737	2 Q6JC13_V9IRU	Q6JC13 adeno-ssoc
7	2832	87.1	736	2 Q6JC08_V9IRU	Q6JC08 adeno-ssoc
8	2831.5	86.8	735	2 Q6JC34_V9IRU	Q6JC34 adeno-ssoc
9	2821.5	86.8	735	2 Q6JC38_V9IRU	Q6JC38 adeno-ssoc
10	2818.5	86.7	735	2 Q6JC21_V9IRU	Q6JC21 adeno-ssoc
11	2818.5	86.7	735	2 Q6JC27_V9IRU	Q6JC27 adeno-ssoc
12	2818.5	86.7	735	2 Q6JC31_V9IRU	Q6JC31 adeno-ssoc
13	2815	86.6	736	2 Q6JC11_V9IRU	Q6JC11 adeno-ssoc
14	2808.5	86.4	735	2 Q6JC17_V9IRU	Q6JC17 adeno-ssoc
15	2807.5	86.4	735	2 Q6JC38_V9IRU	Q6JC38 adeno-ssoc
16	2807.5	86.4	735	2 Q6JC42_V9IRU	Q6JC42 adeno-ssoc
17	2803.5	86.2	735	2 Q6JC34_V9IRU	Q6JC34 adeno-ssoc
18	2802.5	86.2	735	2 Q6JC36_V9IRU	Q6JC36 adeno-ssoc
19	2800	86.1	734	2 Q6JC04_V9IRU	Q6JC04 adeno-ssoc
20	2798	86.1	734	2 Q6JC02_V9IRU	Q6JC02 adeno-ssoc
21	2796.5	86.0	735	2 Q6JC26_V9IRU	Q6JC26 adeno-ssoc
22	2789.5	85.8	735	2 Q6JC43_V9IRU	Q6JC43 adeno-ssoc
23	2785.5	85.7	735	2 Q6JB20_V9IRU	Q6JB20 adeno-ssoc
24	2784.5	85.6	735	2 Q6JB25_V9IRU	Q6JB25 adeno-ssoc
25	2782.5	85.6	735	2 Q6JB24_V9IRU	Q6JB24 adeno-ssoc
26	2781	85.5	734	2 Q6JC03_V9IRU	Q6JC03 adeno-ssoc
27	2778.5	85.5	735	2 Q6JC08_V9IRU	Q6JC08 adeno-ssoc
28	2777.5	85.4	735	2 Q6JC06_V9IRU	Q6JC06 adeno-ssoc
29	2775.5	85.4	735	2 Q6JC08_V9IRU	Q6JC08 adeno-ssoc
30	2775.5	85.4	735	2 Q6JC41_V9IRU	Q6JC41 adeno-ssoc
31	2773.5	85.3	735	2 Q6JC07_V9IRU	Q6JC07 adeno-ssoc

32	2773.5	85.3	735	2 Q6JC29_V9IRU	Q6JC29 adeno-ssoc
33	2771.5	85.3	735	2 Q6JC39_V9IRU	Q6JC39 adeno-ssoc
34	2769.5	85.2	735	2 Q6JC06_V9IRU	Q6JC06 adeno-ssoc
35	2768.5	85.2	735	2 Q6JC25_V9IRU	Q6JC25 adeno-ssoc
36	2766.5	85.1	735	2 Q6JC24_V9IRU	Q6JC24 adeno-ssoc
37	2766.5	85.1	736	2 Q6JC24_V9IRU	Q6JC24 adeno-ssoc
38	2764.5	85.0	735	2 Q6JC52_AAV2	Q6JC52 adeno-ssoc
39	2764.5	85.0	735	2 Q6JC31_V9IRU	Q6JC31 adeno-ssoc
40	2763.5	85.0	735	2 Q6JC01_V9IRU	Q6JC01 adeno-ssoc
41	2763.5	85.0	735	2 Q6JC35_V9IRU	Q6JC35 adeno-ssoc
42	2761.5	84.9	735	2 Q6JC35_V9IRU	Q6JC35 adeno-ssoc
43	2759.5	84.9	598	2 Q6JC53_AAV2	Q6JC53 adeno-ssoc
44	2759.5	84.9	735	2 Q6JB24_V9IRU	Q6JB24 adeno-ssoc
45	2758.5	84.9	735	2 Q6JC26_V9IRU	Q6JC26 adeno-ssoc

ALIGNMENTS

RESULT 1	Q9WB88_V9IRU	PRT;	736 AA.
AC	Q9WB88_V9IRU PRELIMINARY;		
DT	01-NOV-1999 (Tremblrel. 12, Created)		
DT	01-NOV-1999 (Tremblrel. 12, Last sequence update)		
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)		
DE	Capaid protein.		
OS	Adeno-associated virus 1.		
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.		
OX	NCBI_TaxID=85106;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=9921438; PubMed=10196295;		
RA	Xiao W., Chirmule N., Berta S.C., McCullough B., Gao G., Wilson J.M.;		
RT	"Gene therapy vectors based on adeno-associated virus type 1."		
RL	J. Virol. 73:3994-4003(1999).		
DR	EMBL AF063497; AAD2757.1; -; Genomic_DNA.		
DR	SMR; Q9WB88; 217-736.		
DR	GO; GO:0019028; P: viral capsid; IEA.		
DR	GO; GO:0005198; P: structural molecule activity; IEA.		
DR	InterPro; IPR001403; Parvo_coat.		
DR	Pfam; PF00740; Parvo_coat; 1.		
SQ	SEQUENCE 736 AA; 81375 MW; CFABF99BD5CD0595 CRC64;		
Query Match	100.0%; Score 3251; DB 2; Length 736;		
Best local Similarity	100.0%; Pred. No. 5.1e-216;		
Matches	599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 TAPGKKRPVEQSPQEPDSSSGIGTGGQOPAKKRLNCGTGSSEVPDPQLGEPATPPA 60		
DB	138 TAPGKKRPVEQSPQEPDSSSGIGTGGQOPAKKRLNCGTGSSEVPDPQLGEPATPPA 197		
QY	61 VEPITMASGGGAPMADNNEGADGVNAGSNMHCSTWLDGRLVITTTSTRTMALPTVNNHLY 120		
DB	198 VEPITMASGGGAPMADNNEGADGVNAGSNMHCSTWLDGRLVITTTSTRTMALPTVNNHLY 257		
QY	121 KQISSASTGASNDNHYFGYSTWGVYDFENRFCHESPRDWQRLINNMGFRPRGLNFKLP 180		
DB	258 KQISSASTGASNDNHYFGYSTWGVYDFENRFCHESPRDWQRLINNMGFRPRGLNFKLP 317		
QY	181 NIOVKAVTTNDGVTTLANNLTSTVOYFSDSEYOLPVVLASAHQGLPPPADVFMIPQYG 240		
DB	318 NIOVKAVTTNDGVTTLANNLTSTVOYFSDSEYOLPVVLASAHQGLPPPADVFMIPQYG 377		
QY	241 YLTINNGSOAVGRSFFCYLEFPSCMLRTGNFTFSYTEEYVPHSSVYAHSGSLDLRLMP 300		
DB	378 YLTINNGSOAVGRSFFCYLEFPSCMLRTGNFTFSYTEEYVPHSSVYAHSGSLDLRLMP 437		
QY	301 LIDQVLYLNNRTQNGSAQNKDLFLSRGSPAGMSYQPKWMLPGPCYRQORVSKTTDDN 360		
DB	438 LIDQVLYLNNRTQNGSAQNKDLFLSRGSPAGMSYQPKWMLPGPCYRQORVSKTTDDN 497		
QY	361 NSNFTWTGASKYVNLNGBESTINPGTAMASHKDEDEKFFPMGVMIGKSGASGASNTALDN 420		

```
|||||
Db NSNFTWTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 557
498
Qy 421 VMITDEBEIKATNPVATERFCTVAVNFOSSSTDPAIDGVHMGALPGWVMDRDVYLQGP 480
Db 558 VMITDEBEIKATNPVATERFCTVAVNFOSSSTDPAIDGVHMGALPGWVMDRDVYLQGP 617
Qy 481 IMAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTPVANPAPFESATKFAFITQYSTG 540
Db 618 IMAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTPVANPAPFESATKFAFITQYSTG 677
Qy 541 QVSVEIEMELQKENSKRNPPEVQYTSNYAKSANVDFVDDNGLYTEPPIGTRVLTPL 599
Db 678 QVSVEIEMELQKENSKRNPPEVQYTSNYAKSANVDFVDDNGLYTEPPIGTRVLTPL 736
```

RESULT 2

```
06JC08_9VIRU
ID 06JC08_9VIRU PRELIMINARY; PRT; 736 AA.
AC 06JC08;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530611; AAS99296.1; -; Genomic_DNA.
DR SMR; 06JC08; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; P:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81317 MW; 373C7C40975CD9BD CRC64;
```

Query Match 99.4%; Score 3230; DB 2; Length 736;

Best Local Similarity 99.2%; Pred. No. 1.4e-214;

Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSSEVPDQPLGEPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSSEVPDQPLGEPATPAA 197
Qy 61 VGPTTASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHLY 120
Db 198 VGPTTASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHLY 257
Qy 121 KOISSASTGASNDNHYFGYSTPMGWYFDNRFHCHSPDQRLINNMMGFRPKRLNFKLF 180
Db 258 KOISSASTGASNDNHYFGYSTPMGWYFDNRFHCHSPDQRLINNMMGFRPKRLNFKLF 317
Qy 181 NIQVEVTTNDGVTTIANNLSTVOVFSDESYQLPYVLSAHQGLPEFPADVMIPOYG 240
Db 318 NIQVEVTTNDGVTTIANNLSTVOVFSDESYQLPYVLSAHQGLPEFPADVMIPOYG 377
Qy 241 YLTINNGQAVGRSSFYCLEYFPQMLRTGNFTFSYTEEVPHSSYAHQSILDRLAMP 300
Db 378 YLTINNGQAVGRSSFYCLEYFPQMLRTGNFTFSYTEEVPHSSYAHQSILDRLAMP 437
Qy 301 LIDQVLYNRTQNGSGAQNKDILFSRGSFAGMSVQPKNMLPGPCYRQOQVSKTKTDNN 360
Db 438 LIDQVLYNRTQNGSGAQNKDILFSRGSFAGMSVQPKNMLPGPCYRQOQVSKTKTDNN 497
Qy 361 NSNFTWTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 420
```

```
|||||
Db NSNFTWTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 557
498
Qy 421 VMITDEBEIKATNPVATERFCTVAVNFOSSSTDPAIDGVHMGALPGWVMDRDVYLQGP 480
Db 558 VMITDEBEIKATNPVATERFCTVAVNFOSSSTDPAIDGVHMGALPGWVMDRDVYLQGP 617
Qy 481 IMAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTPVANPAPFESATKFAFITQYSTG 540
Db 618 IMAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTPVANPAPFESATKFAFITQYSTG 677
Qy 541 QVSVEIEMELQKENSKRNPPEVQYTSNYAKSANVDFVDDNGLYTEPPIGTRVLTPL 599
Db 678 QVSVEIEMELQKENSKRNPPEVQYTSNYAKSANVDFVDDNGLYTEPPIGTRVLTPL 736
```

RESULT 3

```
06JC12_9VIRU
ID 06JC12_9VIRU PRELIMINARY; PRT; 736 AA.
AC 06JC12;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530607; AAS99292.1; -; Genomic_DNA.
DR SMR; 06JC12; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; P:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81364 MW; F2D75A0662E6376 CRC64;
```

Query Match 99.4%; Score 3230; DB 2; Length 736;

Best Local Similarity 99.5%; Pred. No. 1.4e-214;

Matches 596; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSSEVPDQPLGEPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSSEVPDQPLGEPATPAA 197
Qy 61 VGPTTASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHLY 120
Db 198 VGPTTASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHLY 257
Qy 121 KOISSASTGASNDNHYFGYSTPMGWYFDNRFHCHSPDQRLINNMMGFRPKRLNFKLF 180
Db 258 KOISSASTGASNDNHYFGYSTPMGWYFDNRFHCHSPDQRLINNMMGFRPKRLNFKLF 317
Qy 181 NIQVEVTTNDGVTTIANNLSTVOVFSDESYQLPYVLSAHQGLPEFPADVMIPOYG 240
Db 318 NIQVEVTTNDGVTTIANNLSTVOVFSDESYQLPYVLSAHQGLPEFPADVMIPOYG 377
Qy 241 YLTINNGQAVGRSSFYCLEYFPQMLRTGNFTFSYTEEVPHSSYAHQSILDRLAMP 300
Db 378 YLTINNGQAVGRSSFYCLEYFPQMLRTGNFTFSYTEEVPHSSYAHQSILDRLAMP 437
Qy 301 LIDQVLYNRTQNGSGAQNKDILFSRGSFAGMSVQPKNMLPGPCYRQOQVSKTKTDNN 360
Db 438 LIDQVLYNRTQNGSGAQNKDILFSRGSFAGMSVQPKNMLPGPCYRQOQVSKTKTDNN 497
Qy 361 NSNFTWTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 420
```

```
DB 498 NSNFTWTGASKYLNIGRESIINPGTAMASHKODEDEKFFPMGVMIFGKESASASTALDN 557
QY 421 VMTDEEIEIKATNPVATERFGTVAVNFOSSTDPATGVDHAMGALPGMWODRDVYLOGP 480
DB 558 VMTDEEIEIKATNPVATERFGTVAVNFOSSTDPATGVDHAMGALPGMWODRDVYLOGP 617
QY 481 IWAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAEFSATKFSFITYSTG 540
DB 618 IWAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAEFSATKFSFITYSTG 677
QY 541 QVSVEIEWELOKENSKRNKPEVOYTSNYAKSANDFTVDNGLYTEPRPIGTRYLTRL 599
DB 678 QVSVEIEWELOKENSKRNKPEVOYTSNYAKSANDFTVDNGLYTEPRPIGTRYLTRL 736
```

```
RESULT 4
ID 056137_VIRU PRELIMINARY; PRT; 736 AA.
AC 056137;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus 6.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
RT (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
DR EMBL; AF028704; AAB95450.1; -; Genomic_DNA.
DR SMR; 056137; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;
```

```
Query Match 99.3%; Score 3229; DB 2; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.7e-214;
Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 TAPGKKRPVQESPOEPPSSSGIGTGOQPAKKRLNFGQTGDSSEVDPQPLGSPATPAA 60
DB 138 TAPGKKRPVQESPOEPPSSSGIGTGOQPAKKRLNFGQTGDSSEVDPQPLGSPATPAA 197
QY 61 VGPPTMASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTMALPTYNNHLY 120
DB 198 VGPPTMASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTMALPTYNNHLY 257
QY 121 KOISSASTGASNDNHFGYSTPWGYFDNRFCHFSFRDQRLINNMGFRPKLNFCLF 180
DB 258 KOISSASTGASNDNHFGYSTPWGYFDNRFCHFSFRDQRLINNMGFRPKLNFCLF 317
QY 181 NIOVEKVTINDGVTTIANNLSTYQVFSDEYQLPYVLGSAHOGCLPPPADVFMIPQYG 240
DB 318 NIOVEKVTINDGVTTIANNLSTYQVFSDEYQLPYVLGSAHOGCLPPPADVFMIPQYG 377
QY 241 YLTLLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEEVPHSSVYAHSSQSLDRLLNP 300
DB 378 YLTLLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEEVPHSSVYAHSSQSLDRLLNP 437
QY 301 LIDQYLYLNRTONOGSAGQNKDLLFSRGSPPAGMSVQPKWMLPGPCYRQORVSKTKTDNN 360
DB 438 LIDQYLYLNRTONOGSAGQNKDLLFSRGSPPAGMSVQPKWMLPGPCYRQORVSKTKTDNN 497
QY 361 NSNFTWTGASKYLNIGRESIINPGTAMASHKODEDEKFFPMGVMIFGKESASASTALDN 420
DB 498 NSNFTWTGASKYLNIGRESIINPGTAMASHKODEDEKFFPMGVMIFGKESASASTALDN 557
```

```
QY 421 VMTDEEIEIKATNPVATERFGTVAVNFOSSTDPATGVDHAMGALPGMWODRDVYLOGP 480
DB 558 VMTDEEIEIKATNPVATERFGTVAVNFOSSTDPATGVDHAMGALPGMWODRDVYLOGP 617
QY 481 IWAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAEFSATKFSFITYSTG 540
DB 618 IWAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAEFSATKFSFITYSTG 677
QY 541 QVSVEIEWELOKENSKRNKPEVOYTSNYAKSANDFTVDNGLYTEPRPIGTRYLTRL 599
DB 678 QVSVEIEWELOKENSKRNKPEVOYTSNYAKSANDFTVDNGLYTEPRPIGTRYLTRL 736
```

```
RESULT 5
ID 06JC10_VIRU PRELIMINARY; PRT; 736 AA.
AC 06JC10;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530609; AAS99294.1; -; Genomic_DNA.
DR SMR; 06JC10; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81388 MW; BA9463B970028DF0 CRC64;
```

```
Query Match 98.5%; Score 3203; DB 2; Length 736;
Best Local Similarity 98.8%; Pred. No. 1.1e-212;
Matches 592; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 TAPGKKRPVQESPOEPPSSSGIGTGOQPAKKRLNFGQTGDSSEVDPQPLGSPATPAA 60
DB 138 TAPGKKRPVQESPOEPPSSSGIGTGOQPAKKRLNFGQTGDSSEVDPQPLGSPATPAA 197
QY 61 VGPPTMASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTMALPTYNNHLY 120
DB 198 VGPPTMASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTMALPTYNNHLY 257
QY 121 KOISSASTGASNDNHFGYSTPWGYFDNRFCHFSFRDQRLINNMGFRPKLNFCLF 180
DB 258 KOISSASTGASNDNHFGYSTPWGYFDNRFCHFSFRDQRLINNMGFRPKLNFCLF 317
QY 181 NIOVEKVTINDGVTTIANNLSTYQVFSDEYQLPYVLGSAHOGCLPPPADVFMIPQYG 240
DB 318 NIOVEKVTINDGVTTIANNLSTYQVFSDEYQLPYVLGSAHOGCLPPPADVFMIPQYG 377
QY 241 YLTLLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEEVPHSSVYAHSSQSLDRLLNP 300
DB 378 YLTLLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEEVPHSSVYAHSSQSLDRLLNP 437
QY 301 LIDQYLYLNRTONOGSAGQNKDLLFSRGSPPAGMSVQPKWMLPGPCYRQORVSKTKTDNN 360
DB 438 LIDQYLYLNRTONOGSAGQNKDLLFSRGSPPAGMSVQPKWMLPGPCYRQORVSKTKTDNN 497
QY 361 NSNFTWTGASKYLNIGRESIINPGTAMASHKODEDEKFFPMGVMIFGKESASASTALDN 420
DB 498 NSNFTWTGASKYLNIGRESIINPGTAMASHKODEDEKFFPMGVMIFGKESASASTALDN 557
```

```
QY 421 VMITDEEIKATNPVATERFGTVAVNFQSSSTDPAIDGVHAMGALPGMWODRDVYLQGP 480
DB 558 VMITDEEIKATNPVATERFGTVAVNFQSSSTDPAIDGVHAMGALPGMWODRDVYLQGP 617
QY 481 IWAKIPHTDGHFHPSPMLMGFGGLKNPPOILLIKNTVPANPPAEFSATKFASTITQYSTG 540
DB 618 IWAKIPHTDGHFHPSPMLMGFGGLKNPPOILLIKNTVPANPPAEFSATKFASTITQYSTG 677
QY 541 QVSVEIEIEMELOKENSKRNPPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGRYLRPL 599
DB 678 QVSVEIEIEMELOKENSKRNPPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGRYLRPL 736
```

RESULT 6

```
Q6JC13 9VIRU
ID O6JC13_9VIRU PRELIMINARY; PRT; 737 AA.
AC O6JC13;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530606; AAS9291.1; -; Genomic_DNA.
DR SMR; Q6JC13; 218-737.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvovir.
DR Pfam; PF00740; Parvovir.
SQ SEQUENCE 737 AA; 81332 MW; 94CED0F3AFD68492 CRC64;
```

Query Match 97.3%; Score 3163.5; DB 2; Length 737;
Best Local Similarity 97.3%; Pred. No. 5.7e-210;
Matches 584; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

```
QY 1 TAPGKKRPVEGSPQ-EPDSSSGIGKTGOOPAKKRLNFGQTGDSSEVPDQPIGEPAPTPA 59
DB 138 TAPGKKRPVEGSPQ-EPDSSSGIGKTGOOPAKKRLNFGQTGDSSEVPDQPIGEPAPTPA 197
QY 60 AVGPPTMASGGAPMADNNEGADVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHL 119
DB 198 GAGSGTMAAGGAPMADNNEGADVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHL 257
QY 120 YKOISASTGASNDNHYFGYSTPWGYPDFNRHCHFSPPDMQRLINNMGFRPKLNFYLF 179
DB 258 YKOISASTGASNDNHYFGYSTPWGYPDFNRHCHFSPPDMQRLINNMGFRPKLNFYLF 317
QY 180 FNIOQKEVTTNDGVTIANNLSTVOVFSDESYOLPYVLGSAHQGLPPFPADVFMIPQY 239
DB 318 FNIOQKEVTTNDGVTIANNLSTVOVFSDESYOLPYVLGSAHQGLPPFPADVFMIPQY 377
QY 240 GYTLTNGSOAVGRSSFYCLEYFSPQMLRTGNNTFTSYFEEVPHSSVAHSQSIDRLMNP 299
DB 378 GYTLTNGSOAVGRSSFYCLEYFSPQMLRTGNNTFTSYFEEVPHSSVAHSQSIDRLMNP 437
QY 300 PLIDQYLYLNTQNGSGAQNKDILFSRGSPPAGSVQPKNMLPGPCYRQORVSKTKTDN 359
DB 438 PLIDQYLYLNTQNGSGAQNKDILFSRGSPPAGSVQPKNMLPGPCYRQORVSKTKTDN 497
QY 360 NNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFPFMSGVMI FGKESAGASNTALD 419
DB 498 NNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFPFMSGVMI FGKESAGASNTALD 557
```

```
QY 420 NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPAIDGVHAMGALPGMWODRDVYLQGP 479
DB 558 NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPAIDGVHAMGALPGMWODRDVYLQGP 617
QY 480 PIWAKIPHTDGHFHPSPMLMGFGGLKNPPOILLIKNTVPANPPAEFSATKFASTITQYST 539
DB 618 PIWAKIPHTDGHFHPSPMLMGFGGLKNPPOILLIKNTVPANPPAEFSATKFASTITQYST 677
QY 540 QVSVEIEIEMELOKENSKRNPPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGRYLRPL 599
DB 678 QVSVEIEIEMELOKENSKRNPPEVOYTSNYSKASANDVFTVDNNGLYTEPRPIGRYLRPL 737
```

RESULT 7

```
O56139 9VIRU
ID O56139_9VIRU PRELIMINARY; PRT; 736 AA.
AC O56139;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Capsid protein VPI.
GN Adeno-associated virus 3B.
OS Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68742;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
RT (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
DR EMBL; AF028705; AAB95452.1; -; Genomic_DNA.
DR SMR; O56139; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvovir.
DR Pfam; PF00740; Parvovir.
SQ SEQUENCE 736 AA; 81906 MW; DD52331AD5F0D70F CRC64;
```

Query Match 87.1%; Score 2832; DB 2; Length 736;
Best Local Similarity 85.7%; Pred. No. 4.8e-187;
Matches 514; Conservative 34; Mismatches 50; Indels 2; Gaps 2;

```
QY 1 TAPGKKRPVEGSPQ-EPDSSSGIGKTGOOPAKKRLNFGQTGDSSEVPDQPIGEPAPTPA 60
DB 138 TAPGKKRPVEGSPQ-EPDSSSGIGKTGOOPAKKRLNFGQTGDSSEVPDQPIGEPAPTPA 197
QY 61 VGPTTMAAGGAPMADNNEGADVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHL 120
DB 198 LGSNTMASGGAPMADNNEGADVGNASGNHCDSTWLGDRVITTSRTMALPTYNNHL 257
QY 121 KOISASTGASNDNHYFGYSTPWGYPDFNRHCHFSPPDMQRLINNMGFRPKLNFYLF 180
DB 258 KOISASTGASNDNHYFGYSTPWGYPDFNRHCHFSPPDMQRLINNMGFRPKLNFYLF 316
QY 181 NIQYKEVTTNDGVTIANNLSTVOVFSDESYOLPYVLGSAHQGLPPFPADVFMIPQY 240
DB 317 NIQYKEVTTNDGVTIANNLSTVOVFSDESYOLPYVLGSAHQGLPPFPADVFMIPQY 376
QY 241 YTLTNGSOAVGRSSFYCLEYFSPQMLRTGNNTFTSYFEEVPHSSVAHSQSIDRLMNP 300
DB 377 YTLTNGSOAVGRSSFYCLEYFSPQMLRTGNNTFTSYFEEVPHSSVAHSQSIDRLMNP 436
QY 301 LIDQYLYLNTQNGSGAQNKDILFSRGSPPAGSVQPKNMLPGPCYRQORVSKTKTDN 359
DB 437 LIDQYLYLNTQNGSGAQNKDILFSRGSPPAGSVQPKNMLPGPCYRQORVSKTKTDN 496
QY 360 NNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFPFMSGVMI FGKESAGASNTALD 419
DB 497 NNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFPFMSGVMI FGKESAGASNTALD 556
QY 420 NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPAIDGVHAMGALPGMWODRDVYLQGP 479
```

[illegible]

```

RESULT 8
O67008_9VIRU
ID O67008_9VIRU PRELIMINARY; PRT; 735 AA.
AC O67008;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Capsid protein VP1.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OC NCBI_taxid=272636;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RA Chen C.L., Jensen R.L., Schnepf B.C., Connell M.J., Bartlett J.S.,
RA Sferra T.J., Shell R., Johnson P.R., Clark K.R.;
RT "Characterization of Adeno-Associated Viruses in Children.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY695376; AAU05370.1; -; Genomic_DNA.
DR SMR; O67008; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo Coat.
DR Pfam; PF00740; Parvo_Coat_1.
SQ SEQUENCE 735 AA; 81887 MW; 258BFB955FC925 CRC64;

```

Query Match	87.1%;	Score 2831.5;	DB 2;	Length 735;
Best Local Similarity	85.0%;	Pred. No. 5.1e-187;		
Matches 509;	Conservative 39;	Mismatches 50;	Indels 1;	Gaps 1;

Qy	1	TAPGKKR	PPEOS	POEPDSSSGT	GKTGOOP	PAKKRLNF	QGTGSESE	VPDDPO	PGEPATPA	60	
Db	138	TAPGKKR	PPEHSPAP	DSSSGTGKSGOOP	PARRLNF	QGTGSD	VPDDPO	PGEPAP	PTS	197	
Qy	61	VGPPTMASGGGA	PMADNNEG	ADGVNAGSN	WHMHCDS	TWLGDR	YITVSTR	TWALPT	NNHLY	120	
Db	198	LGSTTMASGGGA	PMVADNNEG	ADGVNS	SCNMHCDS	QWJGDR	YITVSTR	TWALPT	NNHLY	257	
Qy	121	KQISSASTGAS	NDNNH	PGYSTPMWY	PDENR	FHCFS	SPRDMO	RLINNNG	FERKRLNF	180	
Db	258	KQISSOS	-GASNDNNH	PGYSTPMWY	FDENR	FHCFS	PRDMO	RLINNNG	FERKRLNF	316	
Qy	181	NIOVEVLT	NODEV	TIANNLT	STVOVFS	SEVQLPY	VLSAHO	QGLPP	PAPVFM	240	
Db	317	NIOVEVLT	NODEV	TIANNLT	STVOVFS	SEVQLPY	VLSAHO	QGLPP	PAPVFM	376	
Qy	241	YTLTNNGS	QAVR	SSFCLE	FPSPQML	RTGN	FTFSY	TEEV	PFSSV	YAHSSQSLDR	300
Db	377	YTLTNNGS	QAVR	SSFCLE	FPSPQML	RTGN	FTFSY	TEEV	PFSSV	YAHSSQSLDR	436
Qy	301	LDDOYL	YLNRK	ONOGS	QONKDL	FSRBS	PAGMS	QOPK	NKMLPG	PCYROR	360
Db	437	LDDOYL	YLNRK	ONOGS	QONKDL	FSRBS	PAGMS	QOPK	NKMLPG	PCYROR	496
Qy	361	NSNFTW	GASKY	LNLRBS	IINPGT	MASHK	ODEK	FFPM	SGVMT	IFGKES	420
Db	497	NSNFTW	GASKY	LNLRBS	IINPGT	MASHK	ODEK	FFPM	SGVMT	IFGKES	556
Qy	421	WMTDEE	EIKAN	PVATER	PGYAV	YVFGSS	STDP	PAQD	YHANG	ALPGM	480
Db	557	WMTDEE	EIKAN	PVATER	PGYAV	YVFGSS	STDP	PAQD	YHANG	ALPGM	616

Qy	Db
481	617
IMAKIHTGTHPSPSLMGSGFELKNPPOIILIKOTPVANPAPAFSAATKFASTFYOTSG	IMAKIHTGTHPSPSLMGSGFELKNPPOIIMIKOTPVANPPTNFAKFASTFYOTSG
541	677
QSVSELEMELOENSGKRNMPREYQSYNSAKANDPFDNNGLATPERPCTRTPL	QSVSELEMELOENSGKRNMPREYQSYNSAKANDPFDNNGLATPERPCTRTPL
677	735
QSVSELEMELOENSGKRNMPREYQSYNSAKANDPFDNNGLATPERPCTRTPL	QSVSELEMELOENSGKRNMPREYQSYNSAKANDPFDNNGLATPERPCTRTPL

RESULT 9			
ID	Q6JC28_gVIRU	PRELIMINARY;	PRT: 735 AA.
AC	Q6JC28;		
DT	05-JUL-2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Capsid protein VP1.		
GN	Name=capsid		
OS	Adeno-associated virus.		
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.		
OX	NCBI_TaxID=272636;		
RN	[1]		
RP	NCBIOTCOTC SEQUENCE.		
RX	PUMed=15163731; DOI=10.1126/JVI.79.12.6381-6388.2004;		
RA	Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,		
RA	Wilsen J.M.;		
RT	"Clades of adeno-associated viruses are widely disseminated in human		
RT	tissues";		
RL	J. Virol. 78:6381-6388 (2004).		
DR	EMBL; AY530591; AAS9276.1; -; Genomic_DNA.		
DR	SMR; Q6JC28; 217-735.		
DR	GO; GO:0019028; C:Viral capsid; IEA.		
DR	GO; GO:0005198; F:Structural molecule activity; IEA.		
DR	InterPro; IPR001403; Parvo Coat.		
DR	Pfam; PF00740; Parvo Coat; 1.		
SO	SEQUENCE 735 AA; 81897 MM; 197066F8911FAE9B CRC64;		

Query Match	86.8%	Score 2821.5;	DB 2;	Length 735;
Best Local Similarity	-84.5%;	Pred. No. 2.5e-186;		
Matches 506;	Conservative 40;	Mismatches 52;	Indels 1;	Gaps 1

Qy	I	TAPGKKRVRVEOS	POEPPDSSSGIGTKTGOOPAKKRLINFGOTGSESEVDDPOJGCEPATPA	60
Db	138	TAPGKKRVRHVS	PAPPDSSSGTGAKGQOPARKRLINFGOTGDADVPDPOJLQPPAASBG	197
Qy	61	VGPTTMAAGGGA	PMADNNEGADGVGNASGNHCDSTWJGDRVITITSTETWALPTYNNHLY	120
Db	198	LGSTTMAAGSGA	PMADNNEGADGVGNSSGNHCSQWJGDRVITITSTETWALPTYNNHLY	257
Qy	121	KOISSASTGASNDNH	FGYSTPWGYPDNRPHCHFSPPDQORLINNNNGFRKRLNFKLF	180
Db	258	KOISSQS -GASNDNH	FGYSTPWGYPDNRPHCHFSPPDQORLINNNNGFRKRLNFKLF	316
Qy	161	NIQVEEVTNGCVTTI	IANNLTSTVOVFSDEKOLPYVLGSAHQCLPEPPADVFMIPOYG	240
Db	317	NIQVEEVTNQDSTTTI	IANNLTSTVOVFTDSEQLPYVLGSAHQCLPEPPADVFMIPOYG	376
Qy	241	YLTLLNGSOAVGRSS	PYCLEYFPSSOMLRTGNPFESYTFEEVPHSSVAHSOSLDRLAMP	300
Db	377	YLTLLNGSOAVGRS	PYCLEYFPSSOMLRTGNPFESYTFEDVPFHSSVAHSOSLDRLAMP	436
Qy	301	LIDQVLYLNRQONOS	GSAQNKDILFSRGS PAKGMSVOPKMLPGHCYBQOQRVSKTIDTN	360
Db	437	LIDQVLYLNRQONOS	GTLSGRLFSQAGLPNNMSLOAKNMLPGHCYBQOQRVSKOANDNN	496
Qy	361	NSNFWTGA	SKYLNLRGSIINPGTAMASHKODEDKFPFMSGVMI FGKESAGASNTALDN	420
Db	497	NSNFWTATKTHL	NRGDSLVPNPGPAMASHKODEDEKFPFMIGTLLFGKOGTNAANDADLEN	556
Qy	421	WMTTDEELKANNP	ATERFGTVAANPOSSSTDPATGVDHMANGLPGWMOORDVYLOGF	480
Db	557	WMTTDEELRTNPA	TEOYGTVSNLNLOSNSTGPTTGIVNHOGALPGWMOORDVYLOGF	616

```
OY 481 IWAKIPTHGHHFHPPLMGFGGLKNNPPQILIKNTVPANPPAPESATKFASTFOYSTG 540
DB 617 IWAKIPTHGHHFHPPLMGFGGLKNNPPQILIKNTVPANPPPTNFSSAKFASFIQYSTG 676
OY 541 QVSVEIEMELQKENSCKRMNPEVOYTSNTAKSANDVFTVDNNGLYTERPIGTRILTRPL 599
DB 677 QVSVEIEMELQKENSCKRMNPEIQYTSNTAKSVNDVFTVDNNGVSEBRPIGTRILTRNL 735
```

RESULT 10

```
Q6JC34_9VIRU PRELIMINARY; PRT; 735 AA.
ID Q6JC34_9VIRU PRELIMINARY; PRT; 735 AA.
AC Q6JC34;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Capsid protein VPI.
DE Name=cap;
GN Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues."
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530628; AAS99270.1; -; Genomic_DNA.
DR SMR; Q6JC34; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81962 MW; 85DC69C55232D983 CRC64;
```

```
Query Match 86.8%; Score 2820.5; DB 2; Length 735;
Best Local Similarity 84.5%; Pred. No. 3e-186;
Matches 506; Conservative 41; Mismatches 51; Indels 1; Gaps 1;
```

```
OY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKRLNFGQTGDSSEVPDPQIGEPAPPA 60
DB 138 TAPGKKRPVEHSPVDESSSGTGKAGQAPARKRLNFGQTGADSVDPDQIGEPAPPA 197
OY 61 VCPPTMASGGGAPMADNNEGADGVGNAGNWHCSTWLGDRVITSTRTALPTYNHLY 120
DB 198 LGSTTMATGSGAPMADNNEGADGVGNAGNWHCSTWLGDRVITSTRTALPTYNHLY 257
OY 121 KQISSASTGASNDNHYFGYSTPMGYFDPNRPHCHSPRDMQRLINNMGFRPKLNFLKF 180
DB 258 KQISSOS-GASNDNHYFGYSTPMGYFDPNRPHCHSPRDMQRLINNMGFRPKLNFLKF 316
OY 181 NIQKEVTTNDGVTIANNLSTVQVFSDESEYQLPYVLGSAHQGLPEPPADVMIPOYG 240
DB 317 NIQKEVTTNDGVTIANNLSTVQVFSDESEYQLPYVLGSAHQGLPEPPADVMIPOYG 376
OY 241 YLTLLNGSOAVGRSSFYCLEYFPSCQMLRTGNNFTFSYFEEVPHSSVYAHQSGLDRLMP 300
DB 377 YLTLLNGSOAVGRSSFYCLEYFPSCQMLRTGNNFTFSYFEEVPHSSVYAHQSGLDRLMP 436
OY 301 LIDQVLYLNTQNSGSAQKDLFSRGSAPAGMSVOPKMLPGPCYRQQRVSKTKTDNN 360
DB 437 LIDQVLYLNTQNSGSAQKDLFSRGSAPAGMSVOPKMLPGPCYRQQRVSKTKTDNN 496
OY 361 NSNFTWGTASRYNLNGRESIINPGTAMASHKDEDEKFFPMGCVMI FGKESAGASTALDN 420
DB 497 NSNFTWGTATKYHLNGRSLVNPGRPAMASHKDEDEKFFPMHGTILIFGKGTANADADLN 556
OY 421 VMTDEEIRKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDLYLQGP 480
DB 557 VMTDEEIRKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDLYLQGP 616
```

```
OY 481 IWAKIPTHGHHFHPPLMGFGGLKNNPPQILIKNTVPANPPAPESATKFASTFOYSTG 540
DB 617 IWAKIPTHGHHFHPPLMGFGGLKNNPPQILIKNTVPANPPPTNFSSAKFASFIQYSTG 676
OY 541 QVSVEIEMELQKENSCKRMNPEVOYTSNTAKSANDVFTVDNNGLYTERPIGTRILTRPL 599
DB 677 QVSVEIEMELQKENSCKRMNPEIQYTSNTAKSVNDVFTVDNNGVSEBRPIGTRILTRNL 735
```

RESULT 11

```
Q6JB21_9VIRU PRELIMINARY; PRT; 735 AA.
ID Q6JB21_9VIRU PRELIMINARY; PRT; 735 AA.
AC Q6JB21;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Capsid protein VPI.
DE Name=cap;
GN Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues."
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530628; AAS99270.1; -; Genomic_DNA.
DR SMR; Q6JB21; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81868 MW; 1654BD6287F5474 CRC64;
```

```
Query Match 86.7%; Score 2818.5; DB 2; Length 735;
Best Local Similarity 84.5%; Pred. No. 4.1e-186;
Matches 506; Conservative 40; Mismatches 52; Indels 1; Gaps 1;
```

```
OY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKRLNFGQTGDSSEVPDPQIGEPAPPA 60
DB 138 TAPGKKRPVEHSPVDESSSGTGKAGQAPARKRLNFGQTGADSVDPDQIGEPAPPA 197
OY 61 VCPPTMASGGGAPMADNNEGADGVGNAGNWHCSTWLGDRVITSTRTALPTYNHLY 120
DB 198 LGSTTMATGSGAPMADNNEGADGVGNAGNWHCSTWLGDRVITSTRTALPTYNHLY 257
OY 121 KQISSASTGASNDNHYFGYSTPMGYFDPNRPHCHSPRDMQRLINNMGFRPKLNFLKF 180
DB 258 KQISSOS-GASNDNHYFGYSTPMGYFDPNRPHCHSPRDMQRLINNMGFRPKLNFLKF 316
OY 181 NIQKEVTTNDGVTIANNLSTVQVFSDESEYQLPYVLGSAHQGLPEPPADVMIPOYG 240
DB 317 NIQKEVTTNDGVTIANNLSTVQVFSDESEYQLPYVLGSAHQGLPEPPADVMIPOYG 376
OY 241 YLTLLNGSOAVGRSSFYCLEYFPSCQMLRTGNNFTFSYFEEVPHSSVYAHQSGLDRLMP 300
DB 377 YLTLLNGSOAVGRSSFYCLEYFPSCQMLRTGNNFTFSYFEEVPHSSVYAHQSGLDRLMP 436
OY 301 LIDQVLYLNTQNSGSAQKDLFSRGSAPAGMSVOPKMLPGPCYRQQRVSKTKTDNN 360
DB 437 LIDQVLYLNTQNSGSAQKDLFSRGSAPAGMSVOPKMLPGPCYRQQRVSKTKTDNN 496
OY 361 NSNFTWGTASRYNLNGRESIINPGTAMASHKDEDEKFFPMGCVMI FGKESAGASTALDN 420
DB 497 NSNFTWGTATKYHLNGRSLVNPGRPAMASHKDEDEKFFPMHGTILIFGKGTANADADLN 556
OY 421 VMTDEEIRKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDLYLQGP 480
DB 557 VMTDEEIRKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDLYLQGP 616
```


QY 481 IWAKI PHTDGHFHPSPLMGFGSLKNPPQIILIKNTVPANPPAESATYFASFTIOYSTG 540
DB 617 IWAKI PHTDGHFHPSPLMGFGSLKNPPQIIMIKNTVPANPPPTNFSSAKFASFTIOYSTG 676
QY 541 QVSEIEMELOKENSKRNNPEIQTYSNYSKANSVNDFTVNNGLYTEPRPIGRIYLRPL 599
DB 677 QVSEIEMELOKENSKRNNPEIQTYSNYSKANSVNDFTVNNGLYTEPRPIGRIYLRPL 735

RESULT 12

Q6JBZ7_9VIRU PRELIMINARY; PRT: 735 AA.
ID Q6JBZ7_9VIRU
AC Q6JBZ7_9VIRU
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenbergh L.H., Alvir M.R., Lu Y., Calcedo R., Zhou X., Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530622; AAS9307.1; -, Genomic_DNA.
DR SMR; Q6JBZ7; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81947 MW; EBA318E904E0BF0A CRC64;

Query Match 86.7%; Score 2818.5; DB 2; Length 735;
Best Local Similarity 84.3%; Pred. No. 4.1e-186;
Matches 505; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 1 TAGGKRPEVQSPQEPDSSGGIGKTGQOPAKRLNFGQGSSEVPDPQGLGPPATPA 60
DB 138 TAGGKRPEVQSPQEPDSSGGIGKTGQOPAKRLNFGQGSSEVPDPQGLGPPATPA 197
QY 61 VEPPTMASGGAPMADNNEGADGVNAGSNWCHDSQWLDGDRVITTSRTMALPTYNHLY 120
DB 198 VEPPTMASGGAPMADNNEGADGVNAGSNWCHDSQWLDGDRVITTSRTMALPTYNHLY 257
QY 121 KOISSASTGASNDNHFGYSTPWGYFDENRFCHFSRDMORLNNWGRPRKLNFKLF 180
DB 258 KOISSOS -GASNDNHFGYSTPWGYFDENRFCHFSRDMORLNNWGRPRKLNFKLF 316
QY 181 NIQVEKVTNDGVTIANNLSTVOYFSDSEYQLPYLGAHQGLPPPADVFMIPQYG 240
DB 317 NIQVEKVTNDGVTIANNLSTVOYFSDSEYQLPYLGAHQGLPPPADVFMIPQYG 376
QY 241 YLTLLNGSOAVGSSFYCLEYFPPSOMLRGTNNFSTYEEVEVPHSSYAHOSGLDLRLMP 300
DB 377 YLTLLNGSOAVGSSFYCLEYFPPSOMLRGTNNFSTYEEVEVPHSSYAHOSGLDLRLMP 436
QY 301 LIDQVLYLNRTO-NOSGSAQNKDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDN 360
DB 437 LIDQVLYLNRTO-NOSGSAQNKDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDN 496
QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMGSMVIFGKESAGASNTALDN 420
DB 497 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMGSMVIFGKESAGASNTALDN 556
QY 421 NWITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDIAMGALPGMWQODRDVYLQGP 480
DB 557 NWITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDIAMGALPGMWQODRDVYLQGP 616

QY 481 IWAKI PHTDGHFHPSPLMGFGSLKNPPQIILIKNTVPANPPAESATYFASFTIOYSTG 540
DB 617 IWAKI PHTDGHFHPSPLMGFGSLKNPPQIIMIKNTVPANPPPTNFSSAKFASFTIOYSTG 676
QY 541 QVSEIEMELOKENSKRNNPEIQTYSNYSKANSVNDFTVNNGLYTEPRPIGRIYLRPL 599
DB 677 QVSEIEMELOKENSKRNNPEIQTYSNYSKANSVNDFTVNNGLYTEPRPIGRIYLRPL 735

RESULT 13

Q65311_9VIRU PRELIMINARY; PRT: 736 AA.
ID Q65311_9VIRU
AC Q65311_9VIRU
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Capsid protein.
GN Adeno-associated virus 3.
OS Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=46350;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96266430; PubMed=8661429; DOI=10.1006/viro.1996.0367;
RA Muramatsu S., Mizukami H., Young N.S., Brown K.E.;
RT "Nucleotide sequencing and generation of an infectious clone of adeno-associated virus 3.";
RL Virology 221:208-217(1996).
DR EMBL; U48704; AAC55049.1; -, Genomic_DNA.
DR SMR; Q65311; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81660 MW; AFI1E47B5C67A10 CRC64;

Query Match 86.6%; Score 2815; DB 2; Length 736;
Best Local Similarity 85.2%; Pred. No. 7.1e-186;
Matches 511; Conservative 35; Mismatches 52; Indels 2; Gaps 2;

QY 1 TAGGKRPEVQSPQEPDSSGGIGKTGQOPAKRLNFGQGSSEVPDPQGLGPPATPA 60
DB 138 TAGGKGAVDQSPQEPDSSGGIGKSGOPAKRLNFGQGSSEVPDPQGLGPPATPA 197
QY 61 VEPPTMASGGAPMADNNEGADGVNAGSNWCHDSQWLDGDRVITTSRTMALPTYNHLY 120
DB 198 VEPPTMASGGAPMADNNEGADGVNAGSNWCHDSQWLDGDRVITTSRTMALPTYNHLY 257
QY 121 KOISSASTGASNDNHFGYSTPWGYFDENRFCHFSRDMORLNNWGRPRKLNFKLF 180
DB 258 KOISSOS -GASNDNHFGYSTPWGYFDENRFCHFSRDMORLNNWGRPRKLNFKLF 316
QY 181 NIQVEKVTNDGVTIANNLSTVOYFSDSEYQLPYLGAHQGLPPPADVFMIPQYG 240
DB 317 NIQVEKVTNDGVTIANNLSTVOYFSDSEYQLPYLGAHQGLPPPADVFMIPQYG 376
QY 241 YLTLLNGSOAVGSSFYCLEYFPPSOMLRGTNNFSTYEEVEVPHSSYAHOSGLDLRLMP 300
DB 377 YLTLLNGSOAVGSSFYCLEYFPPSOMLRGTNNFSTYEEVEVPHSSYAHOSGLDLRLMP 436
QY 301 LIDQVLYLNRTO-NOSGSAQNKDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDN 360
DB 437 LIDQVLYLNRTO-NOSGSAQNKDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDN 496
QY 360 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMGSMVIFGKESAGASNTALDN 419
DB 497 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMGSMVIFGKESAGASNTALDN 556
QY 420 NWITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDIAMGALPGMWQODRDVYLQGP 479
DB 557 NWITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDIAMGALPGMWQODRDVYLQGP 616

QY 480 PIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTLPVPANPAPAFSATKPFASFTIOYST 539
DB 617 PIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTLPVPANPAPAFSATKPFASFTIOYST 676
QY 540 QVSVIEIEMELQKENSKRWNPEVOYTSNYSKASANDFTVDNNGLYTEBPPIGTRILYTRPL 599
DB 677 QVSVIEIEMELQKENSKRWNPEIOYTSNYSKASVNDFTVDNNGLYSEBPPIGTRILYTRNL 736

RESULT 14

Q6JC17_VIRU
ID Q6JC17_VIRU PRELIMINARY; PRT; 735 AA.
AC Q6JC17;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X., Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human tissues."
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530602; AAS99287.1; -; Genomic_DNA.
DR SMR; Q6JC17; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81887 MW; 709202P5EFFEB6EC CRC64;

Query Match

Best Local Similarity 86.4%; Score 2808.5; DB 2; Length 735;
Matches 505; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 1 TAPGKKRPVEOSPOEBSGGIGKTQQPARKRLNFGCTGSESVDPDPLGEPATPAA 60
DB 138 TAPGKKRPVEHSPVEPDSGGTGKAGQPARKRINCGTGADSVDPDPLGEPAPASG 197
QY 61 VGPPTMAGCGGAPMADNNEGADGVGNAGNHCDSWTWLGDRVITTSRTMALPTYNHLY 120
DB 198 LGSTTMATGSGAPMADNNEGADGVGNAGNHCDSQWLGDRIYITTSRTMALPTYNHLY 257
QY 121 KOISASGASNDNHFGYSTPMGYFDNRFHCHFSPRDMORLINNNGFPRKRLNFKLF 180
DB 258 KOISGOS-GASNDNHFGYSTPMGYFDNRFHCHFSPRDMORLINNNGFPRKRLNFKLF 316
QY 181 NIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLSAHQGLPPPADVFMIPQY 240
DB 317 NIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLSAHQGLPPPADVFMIPQY 376
QY 241 YITLNGSQAAGRSSFYCLEYFPSCMLRTGNFTTSYFEYVPHSSYAHSGSLRLNMP 300
DB 377 YITLNGSQAAGRSSFYCLEYFPSCMLRTGNFTTSYFEYVPHSSYAHSGSLRLNMP 436
QY 301 LIDQVLYLNTQNSGSAONKDLFSRGS PAMGSVOPKMWLPGPCYRQORVSKTKTDNN 360
DB 437 LIDQVLYLNTQNSGTLQOSRLFSQAGPTNMSLOAKMWLPGPCYRQORVSKTDNN 496
QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCVMLFGKESAGASNTALDN 420
DB 497 NSNFTWTATKYLHNGRSLVNPGRPAMASHKDEDEKFFPMHGTLLFGKGTANADADLN 556
QY 421 VMTDEEIRKATNPVATERFGTVAVNFQSSSTDPATGVDYHAMGALPGMYWODRDVYLQCP 480
DB 557 VMTDEEIRATNPVATERFGTVAVNFQSSSTDPATGVDYHAMGALPGMYWODRDVYLQCP 616

QY 481 IMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTLPVPANPAPAFSATKPFASFTIOYSTG 540
DB 617 IMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTLPVPANPAPAFSATKPFASFTIOYSTG 676
QY 541 QVSVIEIEMELQKENSKRWNPEVOYTSNYSKASANDFTVDNNGLYTEBPPIGTRILYTRPL 599
DB 677 QVSVIEIEMELQKENSKRWNPEIOYTSNYSKASVNDFTVDNNGLYSEBPPIGTRILYTRNL 735

RESULT 15

Q6JC38_VIRU
ID Q6JC38_VIRU PRELIMINARY; PRT; 735 AA.
AC Q6JC38;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X., Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human tissues."
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530581; AAS99266.1; -; Genomic_DNA.
DR SMR; Q6JC38; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81913 MW; 58131FDAEB024251 CRC64;

Query Match

Best Local Similarity 86.4%; Score 2807.5; DB 2; Length 735;
Matches 504; Conservative 40; Mismatches 54; Indels 1; Gaps 1;

QY 1 TAPGKKRPVEOSPOEBSGGIGKTQQPARKRLNFGCTGSESVDPDPLGEPATPAA 60
DB 138 TAPGKKRPVEHSPVEPDSGGTGKAGQPARKRINCGTGADSVDPDPLGEPAPASG 197
QY 61 VGPPTMAGCGGAPMADNNEGADGVGNAGNHCDSWTWLGDRVITTSRTMALPTYNHLY 120
DB 198 LGSTTMATGSGAPMADNNEGADGVGNAGNHCDSQWLGDRIYITTSRTMALPTYNHLY 257
QY 121 KOISASGASNDNHFGYSTPMGYFDNRFHCHFSPRDMORLINNNGFPRKRLNFKLF 180
DB 258 KOISGOS-GASNDNHFGYSTPMGYFDNRFHCHFSPRDMORLINNNGFPRKRLNFKLF 316
QY 181 NIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLSAHQGLPPPADVFMIPQY 240
DB 317 NIQVEKVTNDGVTIANNLSTVQVFSDEYQLPYVLSAHQGLPPPADVFMIPQY 376
QY 241 YITLNGSQAAGRSSFYCLEYFPSCMLRTGNFTTSYFEYVPHSSYAHSGSLRLNMP 300
DB 377 YITLNGSQAAGRSSFYCLEYFPSCMLRTGNFTTSYFEYVPHSSYAHSGSLRLNMP 436
QY 301 LIDQVLYLNTQNSGSAONKDLFSRGS PAMGSVOPKMWLPGPCYRQORVSKTKTDNN 360
DB 437 LIDQVLYLNTQNSGTLQOSRLFSQAGPTNMSLOAKMWLPGPCYRQORVSKTDNN 496
QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCVMLFGKESAGASNTALDN 420
DB 497 NSNFTWTATKYLHNGRSLVNPGRPAMASHKDEDEKFFPMHGTLLFGKGTANADADLN 556
QY 421 VMTDEEIRKATNPVATERFGTVAVNFQSSSTDPATGVDYHAMGALPGMYWODRDVYLQCP 480
DB 557 VMTDEEIRITNPVATERFGTVAVNFQSSSTDPATGVDYHAMGALPGMYWODRDVYLQCP 616

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:26:29 ; Search time 27.2418 Seconds
(without alignment)
1817.892 Million cell updates/sec

Title: US-10-696-282-15

Perfect score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPICTRYLTRPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seque, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/iaa/5.COMB.dep: *
2: /cgn2_6/prodata/1/iaa/6.COMB.dep: *
3: /cgn2_6/prodata/1/iaa/H.COMB.dep: *
4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.dep: *
5: /cgn2_6/prodata/1/iaa/RE.COMB.dep: *
6: /cgn2_6/prodata/1/iaa/backfile1.dep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3251	100.0	599	US-09-807-802A-15	Sequence 15, Appl
2	3251	100.0	736	US-09-807-802A-3	Sequence 3, Appl
3	3251	100.0	736	US-09-807-802A-13	Sequence 13, Appl
4	2906	89.4	534	US-09-807-802A-17	Sequence 17, Appl
5	2764.5	85.0	735	US-09-321-589-1	Sequence 1, Appl
6	2764.5	85.0	735	US-10-293-478-1	Sequence 1, Appl
7	2764.5	85.0	735	US-10-038-972A-13	Sequence 13, Appl
8	2759.5	84.9	538	US-10-038-972A-14	Sequence 14, Appl
9	2481.5	76.3	533	US-10-038-972A-15	Sequence 15, Appl
10	1830.5	56.3	598	US-09-532-594B-16	Sequence 16, Appl
11	1830.5	56.3	734	US-09-532-594B-4	Sequence 4, Appl
12	1700.5	52.3	588	US-09-533-427-5	Sequence 5, Appl
13	1700.5	52.3	724	US-09-533-427-4	Sequence 4, Appl
14	1690.5	52.0	544	US-09-532-594B-18	Sequence 18, Appl
15	1665	51.2	532	US-09-533-427-6	Sequence 6, Appl
16	772	23.7	756	US-09-438-268-4	Sequence 4, Appl
17	493	15.2	781	US-10-187-253E-27	Sequence 27, Appl
18	486	14.9	554	US-10-187-253E-29	Sequence 29, Appl
19	486	14.9	554	US-10-187-253E-35	Sequence 35, Appl
20	484	14.9	781	US-10-187-253E-33	Sequence 33, Appl
21	479.5	14.7	543	US-08-856-841-22	Sequence 22, Appl
22	326	10.0	415	US-08-856-841-20	Sequence 20, Appl
23	325	10.0	486	US-08-856-841-19	Sequence 19, Appl
24	325	10.0	500	US-08-856-841-16	Sequence 16, Appl
25	325	10.0	501	US-08-856-841-18	Sequence 18, Appl
26	318	9.8	395	US-08-856-841-13	Sequence 13, Appl
27	310	9.5	264	US-08-856-841-14	Sequence 14, Appl

28	303	9.3	398	2	US-08-856-841-21	Sequence 21, Appl
29	216.5	6.7	579	6	5223424-13	Patent No. 5223424
30	192	5.9	584	2	US-09-022-949-2	Sequence 2, Appl
31	178.5	5.5	387	2	US-08-856-841-17	Sequence 17, Appl
32	132.5	4.1	489	2	US-10-376-397B-4	Sequence 3855, Ap
33	119	3.7	1095	2	US-09-107-532A-3855	Sequence 4, Appl
34	115	3.5	781	2	US-10-164-595-4	Sequence 4, Appl
35	115	3.5	2283	2	US-10-172-502-4	Sequence 30227, A
36	115	3.5	2736	2	US-09-252-991A-10227	Sequence 27, Appl
37	113.5	3.5	655	1	US-08-469-202-27	Sequence 34, Appl
38	113.5	3.5	655	1	US-08-484-434C-34	Sequence 34, Appl
39	113.5	3.5	655	2	US-09-384-361-34	Sequence 14, Appl
40	112.5	3.5	3060	1	US-08-487-826B-14	Sequence 23413, A
41	111.5	3.4	551	2	US-09-248-796A-23413	Sequence 78, Appl
42	110.5	3.4	624	2	US-08-947-965-78	Sequence 28, Appl
43	110.5	3.4	655	1	US-08-469-202-28	Sequence 35, Appl
44	110.5	3.4	655	1	US-08-484-434C-35	Sequence 35, Appl
45	110.5	3.4	655	2	US-09-384-361-35	Sequence 35, Appl

ALIGNMENTS

```
RESULT 1
US-09-807-802A-15
Sequence 15, Application US/09807802A
Patent No. 6759237
GENERAL INFORMATION:
APPLICANT: Xiao, Weidong
TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
FILE REFERENCE: GNVN.031USA
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: US 60/107,114
PRIOR FILING DATE: 1998-11-05
PRIOR APPLICATION NUMBER: PCT/US99/25694
PRIOR FILING DATE: 1999-11-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin version 3.1
SEQ ID NO 15
LENGTH: 599
TYPE: PRT
ORGANISM: AAV-1
US-09-807-802A-15
Query Match 100.0%; Score 3251; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.1e-280; Indels 0; Gaps 0;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TAPGKKRPVEQSPQEPDSSSGIGIKTGQOPAKKRLNFGTGDSESVDPQPLGSPATPAA 60
1 TAPGKKRPVEQSPQEPDSSSGIGIKTGQOPAKKRLNFGTGDSESVDPQPLGSPATPAA 60
61 VEGFTTASGGGAPMANNNEGADGVGASGNWGHCDSTWLGDRVITTSRTTMAIPTNNHL 120
61 VEGFTTASGGGAPMANNNEGADGVGASGNWGHCDSTWLGDRVITTSRTTMAIPTNNHL 120
61 VEGFTTASGGGAPMANNNEGADGVGASGNWGHCDSTWLGDRVITTSRTTMAIPTNNHL 120
121 KOISASTGASNDNHFGYSTPMGYDFPNRPHCFSPRDMQRLINNMGRRPRGLNFKLF 180
121 KOISASTGASNDNHFGYSTPMGYDFPNRPHCFSPRDMQRLINNMGRRPRGLNFKLF 180
121 KOISASTGASNDNHFGYSTPMGYDFPNRPHCFSPRDMQRLINNMGRRPRGLNFKLF 180
181 NIQKEVTNDGVTIANNLSTVQVFSSEYQPLVILGSAHQCLPPPADVFMIPOYG 240
181 NIQKEVTNDGVTIANNLSTVQVFSSEYQPLVILGSAHQCLPPPADVFMIPOYG 240
181 NIQKEVTNDGVTIANNLSTVQVFSSEYQPLVILGSAHQCLPPPADVFMIPOYG 240
241 YLTINNGSQAVGSSSYCLFYPSSQMLRTGNNTFSTYBEVFPFSSYASOSLDRLMP 300
241 YLTINNGSQAVGSSSYCLFYPSSQMLRTGNNTFSTYBEVFPFSSYASOSLDRLMP 300
241 YLTINNGSQAVGSSSYCLFYPSSQMLRTGNNTFSTYBEVFPFSSYASOSLDRLMP 300
301 LIIQVLYLNRITONOGSAQNKDILFSGSGPAGMSVQPKMLGPCYRQORVSKTYTNN 360
301 LIIQVLYLNRITONOGSAQNKDILFSGSGPAGMSVQPKMLGPCYRQORVSKTYTNN 360
```

Db 301 LIDQYLYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQGRVSKTKTDNN 360
Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASTALDN 420
Db 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASTALDN 420
Qy 421 VMTDEEIKATNPATERFGTAVNFOSSSTDPAIGVAMGALPGVMQDRDVTYLOGP 480
Db 421 VMTDEEIKATNPATERFGTAVNFOSSSTDPAIGVAMGALPGVMQDRDVTYLOGP 480
Qy 481 IWAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAATKFSATFYOSTG 540
Db 481 IWAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAATKFSATFYOSTG 540
Qy 541 QVSVEIEMELQENSKRMNPEVOYTSNYSKANSANVDFVNNGLYTEBRPIGTRYLTRPL 599
Db 541 QVSVEIEMELQENSKRMNPEVOYTSNYSKANSANVDFVNNGLYTEBRPIGTRYLTRPL 599

RESULT 2

US-09-807-802A-3
; Sequence 3, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-3

Query Match 100.0%; Score 3251; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.5e-280;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDSESVPPDPLGEPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDSESVPPDPLGEPATPAA 197
Qy 61 VGPPTMASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTWALPTYNNHLY 120
Db 198 VGPPTMASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTWALPTYNNHLY 257
Qy 121 KOISSASTGASNDNHFGYSTPMGYFDENRFCHFSPPDMQRLINNMMGFPRKLNFLKF 180
Db 258 KOISSASTGASNDNHFGYSTPMGYFDENRFCHFSPPDMQRLINNMMGFPRKLNFLKF 317
Qy 181 NIQVEVTNDGVTITANNLSTVOVFSDSSEYOLPYVLGSAHQGLPPFPADVMIPOYG 240
Db 318 NIQVEVTNDGVTITANNLSTVOVFSDSSEYOLPYVLGSAHQGLPPFPADVMIPOYG 377
Qy 241 YLTLLNGSOAVGRSSFYCLEYFSPQMLRTGNFTFSEYVPHSSYAHQSOLDRLMNP 300
Db 378 YLTLLNGSOAVGRSSFYCLEYFSPQMLRTGNFTFSEYVPHSSYAHQSOLDRLMNP 437
Qy 301 LIDQYLYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQGRVSKTKTDNN 360
Db 438 LIDQYLYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQGRVSKTKTDNN 497
Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASTALDN 420

Db 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASTALDN 557
Qy 421 VMTDEEIKATNPATERFGTAVNFOSSSTDPAIGVAMGALPGVMQDRDVTYLOGP 480
Db 558 VMTDEEIKATNPATERFGTAVNFOSSSTDPAIGVAMGALPGVMQDRDVTYLOGP 617
Qy 481 IWAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAATKFSATFYOSTG 540
Db 618 IWAKIPHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAATKFSATFYOSTG 677
Qy 541 QVSVEIEMELQENSKRMNPEVOYTSNYSKANSANVDFVNNGLYTEBRPIGTRYLTRPL 599
Db 678 QVSVEIEMELQENSKRMNPEVOYTSNYSKANSANVDFVNNGLYTEBRPIGTRYLTRPL 736

RESULT 3

US-09-807-802A-13
; Sequence 13, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-13

Query Match 100.0%; Score 3251; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.5e-280;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDSESVPPDPLGEPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDSESVPPDPLGEPATPAA 197
Qy 61 VGPPTMASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTWALPTYNNHLY 120
Db 198 VGPPTMASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTWALPTYNNHLY 257
Qy 121 KOISSASTGASNDNHFGYSTPMGYFDENRFCHFSPPDMQRLINNMMGFPRKLNFLKF 180
Db 258 KOISSASTGASNDNHFGYSTPMGYFDENRFCHFSPPDMQRLINNMMGFPRKLNFLKF 317
Qy 181 NIQVEVTNDGVTITANNLSTVOVFSDSSEYOLPYVLGSAHQGLPPFPADVMIPOYG 240
Db 318 NIQVEVTNDGVTITANNLSTVOVFSDSSEYOLPYVLGSAHQGLPPFPADVMIPOYG 377
Qy 241 YLTLLNGSOAVGRSSFYCLEYFSPQMLRTGNFTFSEYVPHSSYAHQSOLDRLMNP 300
Db 378 YLTLLNGSOAVGRSSFYCLEYFSPQMLRTGNFTFSEYVPHSSYAHQSOLDRLMNP 437
Qy 301 LIDQYLYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQGRVSKTKTDNN 360
Db 438 LIDQYLYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQGRVSKTKTDNN 497
Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASTALDN 420
Db 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASTALDN 557
Qy 421 VMTDEEIKATNPATERFGTAVNFOSSSTDPAIGVAMGALPGVMQDRDVTYLOGP 480

Db 558 VMTTDEEIKATNPVATERFGTVAANFGSSSTDPATGDVHAGALPGMWQODRDVYLQGP 617
QY 481 IWAKIPHTDGHFHPSPLMGFGFKNPPQILIKNTVPANPPAEFSATKFASTTQYSTG 540
Db 618 IWAKIPHTDGHFHPSPLMGFGFKNPPQILIKNTVPANPPAEFSATKFASTTQYSTG 677
QY 541 QVSVEIEMELOKENSKRKNPEVOYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL 599
Db 678 QVSVEIEMELOKENSKRKNPEVOYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL 736

RESULT 4
US-09-807-802A-17
; Sequence 17, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 534
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-17

Query Match 89.4%; Score 2906; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 5e-250;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 MASGGGAPMADNNEGADGVNAGSNHCHDSITWLGDRYITTSRTWALPTYNNHLYKOISS 125
Db 1 MASGGGAPMADNNEGADGVNAGSNHCHDSITWLGDRYITTSRTWALPTYNNHLYKOISS 60
QY 126 ASTGASNDNHFGYSTPFGYFDFNRFCHFSPRDMQRLINNMGFRKRLNFKLFNIQVK 185
Db 61 ASTGASNDNHFGYSTPFGYFDFNRFCHFSPRDMQRLINNMGFRKRLNFKLFNIQVK 120
QY 186 EVTTNDGVTITANNLSTVQVPSDSEYQLPYVLGSAHQGLPPPADVPMIPOYGYLTIN 245
Db 121 EVTTNDGVTITANNLSTVQVPSDSEYQLPYVLGSAHQGLPPPADVPMIPOYGYLTIN 180
QY 246 NSGOAVRSSFYCLEYFSPQMLRTGNNFTFSYEEVFPFSSYAHQSOLDRLMP 305
Db 181 NSGOAVRSSFYCLEYFSPQMLRTGNNFTFSYEEVFPFSSYAHQSOLDRLMP 240
QY 306 LYYLNRTONOGSAGNKKDLFSRGS PAMGSVQPKMWLPGPCYRQORVSKTKTDNNNSFT 365
Db 241 LYYLNRTONOGSAGNKKDLFSRGS PAMGSVQPKMWLPGPCYRQORVSKTKTDNNNSFT 300
QY 366 WTGASKYNLNGRESI INPGTAMASHKDEDEKFPFMSGVMTFGKESAGASNTALDNMTD 425
Db 301 WTGASKYNLNGRESI INPGTAMASHKDEDEKFPFMSGVMTFGKESAGASNTALDNMTD 360
QY 426 EEEIKATNPVATERFGTVAANFGSSSTDPATGDVHAGALPGMWQODRDVYLQGP 485
Db 361 EEEIKATNPVATERFGTVAANFGSSSTDPATGDVHAGALPGMWQODRDVYLQGP 420
QY 486 PHTDGHFHPSPLMGFGFKNPPQILIKNTVPANPPAEFSATKFASTTQYSTG 545
Db 421 PHTDGHFHPSPLMGFGFKNPPQILIKNTVPANPPAEFSATKFASTTQYSTG 480
QY 546 IEMELOKENSKRKNPEVOYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL 599

Db 481 IEMELOKENSKRKNPEVOYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL 534
RESULT 5
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALLI D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match 85.0%; Score 2764.5; DB 2; Length 735;
Best Local Similarity 82.8%; Pred. No. 3.3e-237;
Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

QY 1 TAPGKKRPVEGSPQEPDSSGIGKGGQQAQKRLNFGQGDSESVDDPQPLGEPATPPA 60
Db 138 TAPGKKRPVEGSPQEPDSSGIGKGGQQAQKRLNFGQGDSESVDDPQPLGEPATPPA 197
QY 61 VEPPTMASGGGAPMADNNEGADGVNAGSNHCHDSITWLGDRYITTSRTWALPTYNNHLY 120
Db 198 VEPPTMASGGGAPMADNNEGADGVNAGSNHCHDSITWLGDRYITTSRTWALPTYNNHLY 257
QY 121 KOISASTGASNDNHFGYSTPFGYFDFNRFCHFSPRDMQRLINNMGFRKRLNFKLF 180
Db 258 KOISASTGASNDNHFGYSTPFGYFDFNRFCHFSPRDMQRLINNMGFRKRLNFKLF 316
QY 181 NIQKEVTNDGVTITANNLSTVQVPSDSEYQLPYVLGSAHQGLPPPADVPMIPOYGY 240
Db 317 NIQKEVTNDGVTITANNLSTVQVPSDSEYQLPYVLGSAHQGLPPPADVPMIPOYGY 376
QY 241 YLTNNGSAGVSSFYCLEYFSPQMLRTGNNFTFSYEEVFPFSSYAHQSOLDRLMP 300
Db 377 YLTNNGSAGVSSFYCLEYFSPQMLRTGNNFTFSYEEVFPFSSYAHQSOLDRLMP 436
QY 301 LIDQVLYLNRTONOGSAGNKKDLFSRGS PAMGSVQPKMWLPGPCYRQORVSKTKTDNN 360
Db 437 LIDQVLYLNRTONOGSAGNKKDLFSRGS PAMGSVQPKMWLPGPCYRQORVSKTKTDNN 496
QY 437 LIDQVLYLNRTONOGSAGNKKDLFSRGS PAMGSVQPKMWLPGPCYRQORVSKTKTDNN 496
Db 361 NSNFTWTGASKYNLNGRESI INPGTAMASHKDEDEKFPFMSGVMTFGKESAGASNTALDN 420
QY 497 NSNFTWTGASKYNLNGRESI INPGTAMASHKDEDEKFPFMSGVMTFGKESAGASNTALDN 556
Db 421 VMTTDEEIKATNPVATERFGTVAANFGSSSTDPATGDVHAGALPGMWQODRDVYLQGP 480
QY 557 VMTTDEEIKATNPVATERFGTVAANFGSSSTDPATGDVHAGALPGMWQODRDVYLQGP 616
Db 481 IWAKIPHTDGHFHPSPLMGFGFKNPPQILIKNTVPANPPAEFSATKFASTTQYSTG 540
QY 617 IWAKIPHTDGHFHPSPLMGFGFKNPPQILIKNTVPANPPAEFSATKFASTTQYSTG 676
Db 541 QVSVEIEMELOKENSKRKNPEVOYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL 599
QY 677 QVSVEIEMELOKENSKRKNPEVOYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL 735

RESULT 6
US-10-293-478-1
; Sequence 1, Application US/10293478
; Patent No. 6733757
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALLI D.

```

; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; NUMBER OF SEQ ID NOS: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1
```

```

Query Match      85.0%; Score 2764.5; DB 2; Length 735;
Best Local Similarity 82.8%; Pred. No. 3.3e-237;
Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
```

```

QY      1  TAPGKRPVEOSPOEPPSSSGIGKTGOOPAKKRLNFGQTGDSSEVPPDPOPLGEPPATPAA 60
DB      138  TAPGKRPVEHSPVEPDSSSGTGKAGQOPAKRLNFGQTGADSVPPDPOPLGPPAASG 197
QY      61  VGPPTMASGGA PMADNNEGADGVGNASGNMHCSTWLGDRVITSTRTALPTYNHLY 120
DB      198  LGTNTMATGSGAPMADNNEGADGVGNSSGNMHCSTWMDRVITSTRTALPTYNHLY 257
QY      121  KOISSASTGASNDNHYFGYSTPMGYFDFPNRPHCHFSRDMQRLINNMGFRPKRLNFKLF 180
DB      258  KOISSQS-GASNDNHYFGYSTPMGYFDFPNRPHCHFSRDMQRLINNMGFRPKRLNFKLF 316
QY      181  NIQVEVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPFPADVPMIQYG 240
DB      317  NIQVEVTONDGTITIANLSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVPMIQYG 376
QY      241  YLTNNGSQAVGRSSFYCLEYFPQMLRTGNNFTFSYTEEVPFHSSYAHQSGLDRLMNP 300
DB      377  YLTNNGSQAVGRSSFYCLEYFPQMLRTGNNFTFSYTEEVPFHSSYAHQSGLDRLMNP 436
QY      301  LIDQVLYLNTQONSGSAONKDLFSGSPAGMSVOPKMLPGCYRQGRVSKTKTDNN 360
DB      437  LIDQVLYLSTRTNPSGTTGSRLOFSGAGASDIRDQSRMLPGCYRQGRVSKTKTDNN 496
QY      437  LIDQVLYLSTRTNPSGTTGSRLOFSGAGASDIRDQSRMLPGCYRQGRVSKTKTDNN 496
DB      361  NSNFTWTSKYNLNGRESIINPGTAMASHKDEDEKFFPMGSMIFGKESAGASNTALDN 420
QY      497  NSEYMTGATKYHLNGRDSLVPNGPAMASHKDEDEKFFPMGSMIFGKESAGASNTALDN 556
DB      421  VMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAITGVHAMGALPGMWODRDVYLQGP 480
QY      557  VMTDEBEIRITNPVATEQYGSVNTLQGRNQAATADVNTQGVLPGMWODRDVYLQGP 616
DB      481  IMAKIPHTDGHFSPPLMGSGFGLKNPPOILIKNTPPVNPANPAEFSATKFAFTIYOSTG 540
QY      617  IMAKIPHTDGHFSPPLMGSGFGLKNPPOILIKNTPPVNPANPAEFSATKFAFTIYOSTG 676
DB      541  QVSVEIEMELQKENSKRNPPEIQYTSNKNKSVNVDFTVDITGVSEPRPIGRTYLRNL 599
QY      677  QVSVEIEMELQKENSKRNPPEIQYTSNKNKSVNVDFTVDITGVSEPRPIGRTYLRNL 735
DB
```

```

RESULT 7
US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP1 capsid protien
US-10-038-972A-13
```

```

Query Match      85.0%; Score 2764.5; DB 2; Length 735;
Best Local Similarity 82.8%; Pred. No. 3.3e-237;
Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
```

```

QY      1  TAPGKRPVEOSPOEPPSSSGIGKTGOOPAKKRLNFGQTGDSSEVPPDPOPLGEPPATPAA 60
DB      138  TAPGKRPVEHSPVEPDSSSGTGKAGQOPAKRLNFGQTGADSVPPDPOPLGPPAASG 197
QY      61  VGPPTMASGGA PMADNNEGADGVGNASGNMHCSTWLGDRVITSTRTALPTYNHLY 120
DB      198  LGTNTMATGSGAPMADNNEGADGVGNSSGNMHCSTWMDRVITSTRTALPTYNHLY 257
QY      121  KOISSASTGASNDNHYFGYSTPMGYFDFPNRPHCHFSRDMQRLINNMGFRPKRLNFKLF 180
DB      258  KOISSQS-GASNDNHYFGYSTPMGYFDFPNRPHCHFSRDMQRLINNMGFRPKRLNFKLF 316
QY      181  NIQVEVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPFPADVPMIQYG 240
DB      317  NIQVEVTONDGTITIANLSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVPMIQYG 376
QY      241  YLTNNGSQAVGRSSFYCLEYFPQMLRTGNNFTFSYTEEVPFHSSYAHQSGLDRLMNP 300
DB      377  YLTNNGSQAVGRSSFYCLEYFPQMLRTGNNFTFSYTEEVPFHSSYAHQSGLDRLMNP 436
QY      301  LIDQVLYLNTQONSGSAONKDLFSGSPAGMSVOPKMLPGCYRQGRVSKTKTDNN 360
DB      437  LIDQVLYLSTRTNPSGTTGSRLOFSGAGASDIRDQSRMLPGCYRQGRVSKTKTDNN 496
QY      437  LIDQVLYLSTRTNPSGTTGSRLOFSGAGASDIRDQSRMLPGCYRQGRVSKTKTDNN 496
DB      361  NSNFTWTSKYNLNGRESIINPGTAMASHKDEDEKFFPMGSMIFGKESAGASNTALDN 420
QY      497  NSEYMTGATKYHLNGRDSLVPNGPAMASHKDEDEKFFPMGSMIFGKESAGASNTALDN 556
DB      421  VMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAITGVHAMGALPGMWODRDVYLQGP 480
QY      557  VMTDEBEIRITNPVATEQYGSVNTLQGRNQAATADVNTQGVLPGMWODRDVYLQGP 616
DB      481  IMAKIPHTDGHFSPPLMGSGFGLKNPPOILIKNTPPVNPANPAEFSATKFAFTIYOSTG 540
QY      617  IMAKIPHTDGHFSPPLMGSGFGLKNPPOILIKNTPPVNPANPAEFSATKFAFTIYOSTG 676
DB      541  QVSVEIEMELQKENSKRNPPEIQYTSNKNKSVNVDFTVDITGVSEPRPIGRTYLRNL 599
QY      677  QVSVEIEMELQKENSKRNPPEIQYTSNKNKSVNVDFTVDITGVSEPRPIGRTYLRNL 735
DB
```

```

RESULT 8
US-10-038-972A-14
; Sequence 14, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 14
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP2 capsid protien
US-10-038-972A-14
```

```

Query Match      84.9%; Score 2759.5; DB 2; Length 598;
```


Best Local Similarity 82.8%; Pred. No. 6,6e-237;
Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

QY 2 AAGKRRPVEQSPQEDSSSGIGTKTGOQPAKKRLNFQGTDSBSPVDPPLGEPPTPAAY 61
DB 2 AAGKRRPVEHSPVEDSSSGTGKAGQOPAKKRLNFQGTDSBSPVDPPLGEPPTPAAY 61

QY 62 GPTTAAAGGAPMADNNEGADGVNAGSNMHCDSITWLGDRVITTTSTRTWALPTYNHLYK 121
DB 62 GPTTAAAGGAPMADNNEGADGVNAGSNMHCDSITWLGDRVITTTSTRTWALPTYNHLYK 121

QY 122 QISSASGTASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKLN 181
DB 122 QISSASGTASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKLN 181

QY 182 IQVKEVTMDGVTTANLITSTVOVSDSEYOLPYLGSAHQCLPPFPADVFPMVPOYGY 241
DB 182 IQVKEVTMDGVTTANLITSTVOVSDSEYOLPYLGSAHQCLPPFPADVFPMVPOYGY 241

QY 242 LTLNNGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPL 301
DB 242 LTLNNGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPL 301

QY 302 IQOYLTLNRTONQSSAONKLLFSRGSPPAGMSVOPKXWLPQPCYRQORVSKTKTDNNN 361
DB 302 IQOYLTLNRTONQSSAONKLLFSRGSPPAGMSVOPKXWLPQPCYRQORVSKTKTDNNN 361

QY 362 SNFTWTGASKYNLNGRESIINPCTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNV 421
DB 362 SNFTWTGASKYNLNGRESIINPCTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNV 421

QY 422 MITDEEIKATNPVATERFGTVAANFQSSSTDPATGDVAMGALPGMWMODRDVYLQGP 481
DB 422 MITDEEIKATNPVATERFGTVAANFQSSSTDPATGDVAMGALPGMWMODRDVYLQGP 481

QY 482 WAKIPHTDGHFHPSPLMGFGFLKNPPOLLIKNTVPANPAPFAFSATKTFASFTYOYSTQ 541
DB 482 WAKIPHTDGHFHPSPLMGFGFLKNPPOLLIKNTVPANPAPFAFSATKTFASFTYOYSTQ 541

QY 542 VSVIEIEMELQKENSCKRNPVOYTSNYAKSANDVFTVNNGLYTEBRPIGTRLYTRPL 599
DB 542 VSVIEIEMELQKENSCKRNPVOYTSNYAKSANDVFTVNNGLYTEBRPIGTRLYTRPL 599

QY 599 VSVIEIEMELQKENSCKRNPVOYTSNYAKSANDVFTVNNGLYTEBRPIGTRLYTRPL 599
DB 599 VSVIEIEMELQKENSCKRNPVOYTSNYAKSANDVFTVNNGLYTEBRPIGTRLYTRPL 599

RESULT 9
US-10-038-972A-15
Sequence 15, Application US/10038972A
Patent No. 6962815
GENERAL INFORMATION:
APPLICANT: J. Bartlett
TITLE OF INVENTION: AAV VECTORS AND METHODS
FILE REFERENCE: 28335/3696US
CURRENT APPLICATION NUMBER: US/10/038, 972A
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: US 60/260,124
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 533
TYPE: PRT
ORGANISM: adeno-associated virus 2 VP3 capsid protein
US-10-038-972A-15

Query Match 76.3%; Score 2481.5; DB 2; Length 533;
Best Local Similarity 83.3%; Pred. No. 3.2e-212;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 66 MASGGAPMADNNEGADGVNAGSNMHCDSITWLGDRVITTTSTRTWALPTYNHLYK 125
DB 66 MASGGAPMADNNEGADGVNAGSNMHCDSITWLGDRVITTTSTRTWALPTYNHLYK 125

QY 120 YKOISSASGTASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKL 179
DB 120 YKOISSASGTASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKL 179

QY 126 ASTGASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKLNIQVK 185
DB 126 ASTGASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKLNIQVK 185

DB 61 QS-GASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKLNIQVK 119
QY 186 EVTTNDGVTTANLITSTVOVSDSEYOLPYLGSAHQCLPPFPADVFPMVPOYGYTLN 245
DB 120 EVTTNDGVTTANLITSTVOVSDSEYOLPYLGSAHQCLPPFPADVFPMVPOYGYTLN 179

QY 246 NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPLIOY 305
DB 246 NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRMLNPLIOY 305

QY 306 LYLNRTONQSSAONKLLFSRGSPPAGMSVOPKXWLPQPCYRQORVSKTKTDNNNSFT 365
DB 306 LYLNRTONQSSAONKLLFSRGSPPAGMSVOPKXWLPQPCYRQORVSKTKTDNNNSFT 365

QY 366 WTGASKYNLNGRESIINPCTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 425
DB 366 WTGASKYNLNGRESIINPCTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 425

QY 426 EEEIKATNPVATERFGTVAANFQSSSTDPATGDVAMGALPGMWMODRDVYLQGPWAKI 485
DB 426 EEEIKATNPVATERFGTVAANFQSSSTDPATGDVAMGALPGMWMODRDVYLQGPWAKI 485

QY 486 PHTDGHFHPSPLMGFGFLKNPPOLLIKNTVPANPAPFAFSATKTFASFTYOYSTQVSE 545
DB 486 PHTDGHFHPSPLMGFGFLKNPPOLLIKNTVPANPAPFAFSATKTFASFTYOYSTQVSE 545

QY 546 IEMELQKENSCKRNPVOYTSNYAKSANDVFTVNNGLYTEBRPIGTRLYTRPL 599
DB 546 IEMELQKENSCKRNPVOYTSNYAKSANDVFTVNNGLYTEBRPIGTRLYTRPL 599

RESULT 10
US-09-532-594B-16
Sequence 16, Application US/09532594B
Patent No. 6468524
GENERAL INFORMATION:
APPLICANT: Chorini, John A.
APPLICANT: Kotin, Robert M.
APPLICANT: Saefer, Brian
APPLICANT: Davidson, Beverly
TITLE OF INVENTION: AAV VECTORS AND USES THEREOF
FILE REFERENCE: 14014.0252U2
CURRENT APPLICATION NUMBER: US/09/532,594B
CURRENT FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 598
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
NAME/KEY: misc_feature
OTHER INFORMATION: AAV4 capsid protein VP2
US-09-532-594B-16

Query Match 56.3%; Score 1830.5; DB 2; Length 598;
Best Local Similarity 57.8%; Pred. No. 3.7e-154;
Matches 355; Conservative 74; Mismatches 154; Indels 31; Gaps 10;

QY 1 TAPGKRRPVEQSPQEDSSSGIGTKTGOQPAKKRLNF-GGTGSESVDPPLGEPPTPA 59
DB 1 TAPGKRRPVEQSPQEDSSSGIGTKTGOQPAKKRLNF-GGTGSESVDPPLGEPPTPA 59

QY 60 AVPTTMAAGGAPMADNNEGADGVNAGSNMHCDSITWLGDRVITTTSTRTWALPTYNHLYK 119
DB 60 AVPTTMAAGGAPMADNNEGADGVNAGSNMHCDSITWLGDRVITTTSTRTWALPTYNHLYK 119

QY 120 YKOISSASGTASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKL 179
DB 120 YKOISSASGTASNDNHFGYSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKL 179

QY 115 YKRLGE-----SLOSTNYNGFSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKL 170
DB 115 YKRLGE-----SLOSTNYNGFSTPMGYFPFNRPHCHFSRDMQRLINNMWGRPRKRLNFKL 170

```

Qy 180 FNIOQWERTTNDGVTIANNILSTVOYPSDSBYOLPYTLGSAGHGLRPPRPADVFMIPQY 239
Db 171 FNIOQWERTTNSNETTVANNILSTVOIFADSSYELPYVMADAGQSGSLRPPRPADVFMIPQY 230
Qy 240 GY---LTLNNGSQAVRSFSFYCLFEYFPSSOMLRTGNNFTSYFEEVPHSHYASQSLDR 296
Db 231 GYGGVLVTGNTSQOQDTDRNAFYCLFEYFPSSOMLRTGNFNFTYSFEXVFPFHMYAHQSGLDR 290
Qy 297 LNNPLIDQYLYNFRQN---QSSAQKDLIFSRGSPAGMSYQPKMYLPGPYRQORV 352
Db 291 LNNPLIDQYLMGLOSTTTGTTLNAGATTN--FPKJLRPTNFSPPKKNMLRGPRIKQGF 347
Qy 353 SKTKTNNNSNFTWTAS---KY---NLNGRESIINPGTMAASHKDEDEKFFMSSGVMI 405
Db 348 SKTA--NOWNKIPATSDSLIKYETHSLTDGMSMLTGPMPMATAGRADSK--PBNQSLIR 404
Qy 406 FGKESAGASNTALDNMTDSEBEIKATNEVATERFQVAVNFPQSSSTDPATGDVHAMGAL 465
Db 405 AGCRKONGNTATVPGLTIFPTESEELATYATDTDMGNLPGCDQGSNSNLPYDURLTALGAV 464
Qy 466 PGWVMODRDVYLOGPIMAKIPHTDGHFHSPLMGSGGLKNRPQILIKNTPVNPANPAEF 525
Db 465 PGWVMODRDVYVGGPIMAKIPHTDGHFHSPLIGSPGLKHPPOPIFNTPVNPANPATTF 524
Qy 526 SATKFASTFOYSTGQVSEIEMELOKENSKRWNEVOYTSNVAKSANDVFTYDNNGLYT 585
Db 525 SSTPVNSFIQYSTGQVSOQIMEIQKERSKRWNEVOYFTSNYQGNQSLMLAPDAKXYT 584
Qy 586 EPPRIGTRITRPL 599
Db 585 EPPRIGTRITRTHL 598

```

```

RESULT 11
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kottin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc_feature
; OTHER INFORMATION: AAV4 capsid protein VP1
US-09-532-594B-4

```

Query Match	56.3%	Score 1830.5	DB 2	Length 734
Best Local Similarity	57.8%	Pred. No.5.2e-154		
Matches	355	Conservative 74	Mismatches 154	Indels 31
			Gaps	10
OY	1	TAPGKKRPVPSROEDSSSGIGTKGGQPAKRLNF-GOTGDSESVDPQPLGEPPAPRA	59	
	:			
Dd	137	TAPGKKRPIRESPOQPDSDSTGIGKKYGKQPAKKLLFEEDTAGDGPPEGSTG-----A	190	
OY	60	AVGPTTMSGGCAPADANNRGADGVGNASGMNHCDSITVJGRVITTSRTALPTYNNHL	119	
	:			
Dd	191	MSDDEMAAAAGCAVBEQGQADGVGNASGDMHCDSTISBGIYTTSIRTVLPFYNNHL	250	
OY	120	YKQISSASTGASNDNHRYFGYSTPMWGYFDENRFCHFSPPDWORLINNMGFAPKRLNFKL	179	

```

Db      251  YKRJGE---SLOSNTYNGFSTPMKYPDPENRCHCFSPEDMORRLINNNGMKPKMARVXI 306
QY      180  FNIQVEVTTNDGVTTIANNLSTIVQVSDSEYQLPYVLGSAHQGLPPPADVEMIPOY 239
Db      307  FNIQVEVTTNSGETTVANNLSTVQIFADSSYELPYVMDAQOEGSLPPFPMDVFMVPOY 366
QY      240  QY---LTIANGSAVGRSSFYCLEFFPBPOMLTGNNFFTSYFEEVPEHFSVAHQSGLSLR 296
Db      367  GYCGVLVNTGTSQQQOTDRNAFYCLEFFPBPOMLTGNNFELTYSPEKYPFHSMVAHSQSLDR 426
QY      297  LNNPLIIDQYLYLVNLTQN---QSGSAONKDLLFPRGSPAGASVOVKMMLPGPCYQOQHV 352
Db      427  LNNPLIIDQYLVNLTGLOSTTGTTLNACTATLN---FTKLRPTNSNFPKQMLPGPSIQOGE 483
QY      353  SKTKTDNNNSNFTWTGAS---KY---NINGRESIINPGTAMASHKODEKCFPMGCVMI 405
Db      484  SKTA--NQNYKIPATGSDSLIKYETHSTLDGRMSALTPQPPMATAGPADSK--FSNQLITF 540
QY      406  FGKESAGASNTALDNVMTTDEBEIKATNPVATERGTVAVNVOSSSTDPATGDAVMAGAL 465
Db      541  AGPKONGNTATVPGLTITSEBELAATATDIDMGNLPGQDOSNSNLTPTVRLTALGAN 600
QY      466  PGMWQODRVDYVLOGGIMAKIPTDGHFSPPLMGGGGLGNPPQIILKNTPYVPANPAEF 525
Db      601  PGMWQONDYIYQGIIMAKIPTDGHFSPPLMGGGGLGNPPQIILKNTPYVPANPAATTF 660
QY      526  SATKFASEITQYSTQVSEIEMELOKENSKEKNPEVOYTSNYANSANDFTVNNGLYT 585
Db      661  SSTPNSTFQYSTQVSDVQIDMEIQKERSKKNPEVQTSNYNGQONSLLMAPDAAGKYT 720
QY      586  EPRPIGTRYLRPL 599
Db      721  EPRRIGTRYLRTHL 734

```

```

RESULT 12
US-09-533-427-5
/ Sequence 5, Application US/09533427
/ Patent No. 6855314
/ GENERAL INFORMATION:
/ APPLICANT: Chiorini, John
/ APPLICANT: Kotin, Robert M.
/ APPLICANT: Safer, Brian
/ APPLICANT: Davidson, Elizabeth
/ APPLICANT: Zahner, Joseph
/ TITLE OF INVENTION: AAV5 VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
/ FILE REFERENCE: 14014.0323U2
/ CURRENT APPLICATION NUMBER: US/09/533,427
/ CURRENT FILING DATE: 2000-03-22
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5
/ LENGTH: 588
/ TYPE: PRY
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
/ OTHER INFORMATION: synthetic construct
US-09-533-427-5

```

	Query Match	52.3%	Score 1700.5;	DB 2;	Length 588;
	Best Local Similarity	55.0%;	Pred. No. 1.4e-142;		
	Matches 332;	Conservative 64;	Mismatches 173;	Indels 35;	Gaps 7
QY	3	PEKKRPVQSQRPEPSSSGIGKTGGQPAKKNLNFQTGDSSESVPPQLGEPAPPA-AV	61		
		:::::	:::::	:::::	:::::
Db	13	PKRKARTEEDSKSTSS-----DAEGPGSQQOLQIPAQPASSL	52		
QY	62	GPTTAAAGCGAPMADNNEGADVGNASGNWHCDSTWLDGVITTSRTMALPTYNNHL-YK	121		
		:::::	:::::	:::::	:::::
Db	53	GADVTASAGGQPLGNNNGADVGNASGDHMCDSITWMDRIVATYSTRTWVLPSYNNHQYR	112		

OY	122	QISSASTSANDNHFGVSTPMGWYDPMRHFHGFSPRDMQRLINNNMGPRPRLNFKLEN	181
Db	113	EIKSGVSGSNAANAIFYGISTPMGWYDFNFRHSHMSPRDMQRLINNYMGPRPSLRKIN	172
OY	182	IQVKEVTTNDGVTTIANNLTSFVQVFSDEYOLPYVLGSAMOGCLPPFPADVFMIPQYGY	241
Db	173	IQVKEVTTQDSPTTIANNLTSFVQVFTDDQDLPYVGNSTEGCLPAFPPOVFTLPQYGY	232
OY	242	LTLN--NGSQAVGRSSFYCLEYFPSCOMLRTGNFTFSTYFEEVYFPHSSIAHQSOLDRLN	299
Db	233	ATLNDNTENPERSSPFCELEYFPSCMLRTGNFETYNFEEVYFPFSSAPSONLFXKLN	292
OY	300	PLIDPYLYYINRTQVQSGSAONKDLFSRGSBAGMSVQ--KNMLPPCYRROORVSKTKTD	358
Db	293	PLVDDYLYRFPVSTNNTGVOFQFNKL-----AGRYYANTYKMPFQPMGRTOQMNIGSGV	345
OY	359	NNNSNFTWGSASKYMLNGRESIIINEGTAWASHKDEDEKFFPMGSWIMPEKESAGASNTLY	418
Db	346	NRASISAFATTNRMLEEGASYOVPPQPNQMTNNLQGSNTYALENTMIFMSQPNPCTAT	405
OY	419	---DNVMTTDEBEIKATNPVATERFGTVAVNFQSSSTDPAEDVHAMGALPGHWODRDV	475
Db	406	YLEGMMLITSESETQPVNRVAVYVNGQMATNNQSSSTTAPATGTYNLOEITVPGSWMERDV	465
OY	476	YLQGIYMAKIPPTDGHFHPSPMLGCGGLKNPPQOLIKNTPVPANPPAEFSAKFKSFTT	535
Db	466	YLQGIYMAKIPETGAHFHPSPAMGCGGLGHPMMELIKNTPVPGN--ITGFSDPVSSFTT	524
OY	536	OYSTGOVSVEIEMELQKENSKRMPNVOYTSNPAKSANDVETVDDNNGLTPERPICGRYL	595
Db	525	OYSTGOVTVEMEWELKENSKRMPNVOYTNNNDPQVDFPADPSIGERKTRTPIGIRYL	584
OY	596	TRPL 599	
Db	585	TRPL 588	

```

RESULT 13
US-09-533-427-4
; Sequence 4, Application US/09533427
; Patent No. 6855314
; GENERAL INFORMATION:
; APPLICANT: Chiorini, John
; APPLICANT: Kotin, Robert M.
; APPLICANT: Saefer, Brian
; APPLICANT: Davidson, Elizabeth
; APPLICANT: Zahner, Joseph
; TITLE OF INVENTION: NAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
; FILE REFERENCE: 14014.032302
; CURRENT APPLICATION NUMBER: US/09/533,427
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: /No. 6855314e =
US-09-533-427-4

```

```

Query Match 151; Score 1700.5; DB 2; Length 724;
Best local Similarity 55.2%; Pred. No. 1.9e-142;
Matches 332; Conservative 64; Mismatches 173; Indels 35; Gaps 7;

QY      3 PCKKRVVGGSPGEPSSSGIGKTGGQAAPKKRINFGQTDSSEVPNPGGLGEPTATPA-AV 61
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db       149 PKKKRKRTEDSKFTST-----DAEKGSGSQQLQIPAPASL 188

QY      62 GTTMAAGCGCAPADNNEGADVGNASGNWCHDSTWLGDRIYTITSTRMALPTYNNHLKY 121
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db       189 GADTMSAGGCGGLGDNNOCADVGNGASGNWHCDSTWMGDRIYVTKSTRTWLPSTYNHDIYR 248

```

QY	122	QISSASGASNDNHYFGSTPMGJEDFNRPFCHFSFPRDMORLJNNMGRPRRLNFKLN	181
Db	249	EIKSGSDGSIANAYFGISTPMGJEDFNRPFCHFSHMSPRDMORLJNNMGRPRSLRKLTN	308
QY	182	IQVKEVTTNDGVTIANNLTSTVOYFSDSEYQLPYVLGSAHQCLPPFPADYFMIDYGY	241
Db	309	IQVEEVYQDSTTTIANNLSTVOYFTDDYQLPYVNGTGECLPAFPPOVFTLPOYGY	368
QY	242	LTIN-NGSQAVGSSPYCLEYFBSQMTRTGNFFSTFEVPPFTHSSYAHOSQLDRLN	299
Db	369	ATLRNDTENPTESSPFCEYFBSKMTRTGNFFFTYFBEVPPHSSAPAPQNLFKLN	428
QY	300	PLIIOYLYLNRKTQNGSSAONKDLFBSRGPAGMSVQP-KMWLGPCYCRORRVSXTKD	358
Db	429	PLVIOYLYFVSTNNGTQVFPKLN-----AGRYANTYKMPFEPQWRTOGMNIGSV	481
QY	359	NNNSFTWTGASKYVLNGRESITNGTAMASHKDEDEKFFPMGSVMIPEKESAGASNTAL	418
Db	482	NRASVASATNNRMBEGASYOVPQPPQMGMTNNLOGSNLYALENTMIFFISQAPNPGTTAT	541
QY	419	--DNVMTDEEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGKMWODRDY	475
Db	542	YLEGMLTTSSETPPVNRVAVYNGQWATNNQSSSTAPATGTYNLOEIVPGSVMMERDV	601
QY	476	YLOGPIWAKIPHTDGHFHPSPLMGFGKLNPPQILIKTVPYPAHPAAFSATKFSFT	535
Db	602	YLOGPIWAKIPETGAFHFPSPMGFGFKHPMMILIKTVPBGN-ITSFSDVPVSSFTT	660
QY	536	OYSTGQVSEVELEMELOKENSKRMPNEVOYQTSNAYASANDPFTVDDNGLTEPPICSTRYL	595
Db	661	OYSTGQVSEVELEMELEKENSKRMPNEILOYTNNYNDPQFVDEAPDSTGSEYRTTRPIGSTRYL	720
QY	596	TRPL 599	
Db	721	TRPL 724	

```

US-09-532-594B-18
; Sequence 18, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Corlin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 544
; TYPE: prt
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP3
US-09-532-594B-18

```

[illegible]

```

Db      67  ---SIQSITVNGFSPFMGYPDENRFHCHFPSRDMQRLINNMMGMRPKAMRYKIFNIQYKE 123
Qy      187 VTTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFMIPQYGT---LT 243
Db      124 VTTSGEFTVANNLSTVQIFADSSYELPYVMDAQEGSLPPFPADVFMIPQYCGLYT 183
Qy      244 LUNGSAVGRSSPYCLEYFPSPOMLRTGNNTFFSYFEEVPPFSSVAHSQSLDRLMNPILID 303
Db      184 GNTSQOQTDRAAFYCLEYFPSPOMLRTGNNTFFSYFEEVPPFSSVAHSQSLDRLMNPILID 243
Qy      304 QYLYLNRTON----QSGSAQNKDILFSGSPAGMSVOPKMWLPGPCYRQORVSKTKTDN 359
Db      244 QYLMWLGSTTGTTLNAGATATN---FTKLRPTNSNFKOMWLPSPSKQOGFSKTA--N 298
Qy      360 NNSNTTWGAS---KY----NLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAG 412
Db      299 QNYKIPATGSDSLIREYTHSTLDGWSALTPGPATAGPADSK--FNSNOLIFAGPKONG 357
Qy      413 ASNTALDNVMTDEBEIKATNPATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQD 472
Db      358 NTAATPGULIFTESEELATNATDTDMGONLPGGQSNLFTVRLTLALGAVPGMWQON 417
Qy      473 RDVYLQGPIMAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTVPANPAPAFSATKPEAS 532
Db      418 RDIYQGPIMAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTVPANPAPAFSATKPEAS 477
Qy      533 FITOSTQGVSEIEMWLOKENSKRMPNPEVOYTSNYAKSANDFTVDNNGLYTEPRPIGT 592
Db      478 FITOSTQGVSEIEMWLOKENSKRMPNPEVOYTSNYAKSANDFTVDNNGLYTEPRPIGT 537
Qy      593 RYLTRPL 599
Db      538 RYLTHHL 544

RESULT 15
US-09-533-427-6
; Sequence 6, Application US/09533427
; Patent No. 6855314
; GENERAL INFORMATION:
; APPLICANT: Chlorin, John
; APPLICANT: Kolin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Elizabeth
; APPLICANT: Zahner, Joseph
; TITLE OF INVENTION: AAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
; FILE REFERENCE: 14014.032302
; CURRENT APPLICATION NUMBER: US/09/533,427
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
US-09-533-427-6

```

```

Query Match      51.2%; Score 1665; DB 2; Length 532;
Best Local Similarity 58.9%; Pred. No. 17e-139;
Matches 318; Conservative 54; Mismatches 154; Indels 14; Gaps 5;

```

```

Qy      66 MASGGAPVADNNEGADGVGNAAGNHCDSITWLGDRVITTSRTTALPTYNHLYKOISS 125
Db      1  MASGGGGPLGDNNQAGDGVGNAAGNHCDSITWLGDRVITTSRTTALPTYNHLYKOISS 60
Qy      126 ASTGASNDNHYTGYSTPMGYFPDPRNFHCHFPSRDMQRLINNMMGMRPKAMRYKIFNIQYK 185
Db      61 GSVDSGNANAYFGYSTPMGYFPDPRNFHCHFPSRDMQRLINNMMGMRPKAMRYKIFNIQYK 120

```

```

Qy      186 EVTTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFMIPQYGLTLN 245
Db      121 EVTVQDSTTTIANNLSTVQVFTDDYQLPYVNGNGEGCLPAPPPQVFTLPQGYATLN 180
Qy      246 --NSQA VGRSSPYCLEYFPSPOMLRTGNNTFFSYFEEVPPFSSVAHSQSLDRLMNPILID 303
Db      181 RDNTENPTERSSPFCLEYPSPOMLRTGNNTFFSYFEEVPPFSSVAHSQSLDRLMNPILID 240
Qy      304 QYLYLNRTONQSGSAQNKDILFSGSPAGMSVOP--KNWLPGPCYRQORVSKTKTDNNS 362
Db      241 QYLRFPVSTNNTGVQENKNL-----AGRYANTYKMFPGPMGRTOGNLGSQVNRAS 293
Qy      363 NFWTQASKYVNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTAL---D 419
Db      294 VSAFATINRMELBEGASVQVPPQPMGNTNNGSNTALNTMINSQPANGTATYLEG 353
Qy      420 NMVITDEBEIKATNPATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDQDYLQ 479
Db      354 NMLITSESETOPVNRVAVNVGGMATNQSSTTAPATGYNLQELVGSVMMEKDYVLOG 413
Qy      480 PIMAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTVPANPAPAFSATKPEASFTIOYST 539
Db      414 PIMAKIPETGAHFHSPAMGGFGGLKNPPQILIKNTVPANPAPAFSATKPEASFTIOYST 472
Qy      540 GQVSEIEMWLOKENSKRMPNPEVOYTSNYAKSANDFTVDNNGLYTEPRPIGT 599
Db      473 GQVTVEMWELKENSKRMPNPEVOYTSNYAKSANDFTVDNNGLYTEPRPIGT 532

```

```

Search completed: November 23, 2005, 17:42:48
Job time : 29.2418 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:24:49 ; Search time 95.4286 Seconds
(without alignments)
2458.679 Million cell updates/sec

Title: US-10-696-282-17

Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPICTRYLTRPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2906	100.0	534	3	AAV71169
2	2906	100.0	599	3	AAV71168
3	2906	100.0	736	3	AAV71167
4	2906	100.0	736	6	ABR80232
5	2906	100.0	736	7	ABR62762
6	2906	100.0	736	7	ADP76565
7	2906	100.0	736	9	ADV70294
8	2906	100.0	736	9	ADZ27069
9	2906	100.0	736	9	ADZ27012
10	2902	99.9	735	9	ADV67506
11	2892	99.5	736	9	ADZ27008
12	2892	99.5	737	9	ADZ27086
13	2888	99.4	736	9	ADZ27074
14	2885	99.3	736	4	AAV71167
15	2884	99.2	736	4	AAV71169
16	2884	99.2	736	7	ADP76565
17	2884	99.2	736	7	ADV70293
18	2884	99.2	736	9	ADZ27070
19	2883	99.2	737	9	ADZ27010
20	2863	98.5	736	9	ADZ27009
21	2532.5	87.3	735	9	ADZ26993
22	2532.5	87.1	735	9	ADZ26996
23	2530.5	87.1	735	9	ADZ27034
24	2530	87.1	736	4	AAV59846

25	2529.5	87.0	735	9	ADZ27001	Adz27001 Adeno-sss
26	2529.5	87.0	735	9	ADZ27076	Adz27076 Adeno-sss
27	2529.5	87.0	735	9	ADZ27000	Adz27000 Adeno-sss
28	2528	87.0	736	4	AAV59845	AAV59845 AAV3A cap
29	2528	87.0	736	6	ABR80233	ABR80233 AAV3A vpl
30	2528	87.0	736	7	ABR62763	ABR62763 Adeno-sss
31	2528	87.0	736	7	ADP76572	ADP76572 Adeno-sss
32	2528	87.0	736	7	ADV67507	ADV67507 Amino aci
33	2528	87.0	736	9	ADZ27067	Adz27067 Adeno-sss
34	2527.5	87.0	735	9	ADZ26992	Adz26992 Adeno-sss
35	2524.5	86.9	735	9	ADZ27003	Adz27003 Adeno-sss
36	2524.5	86.9	735	9	ADZ26998	Adz26998 Adeno-sss
37	2522.5	86.8	735	9	ADZ27002	Adz27002 Adeno-sss
38	2520.5	86.7	735	9	ADZ26991	Adz26991 Adeno-sss
39	2520.5	86.7	735	9	ADZ27004	Adz27004 Adeno-sss
40	2517	86.6	734	9	ADZ27036	Adz27036 Adeno-sss
41	2515.5	86.6	735	9	ADZ26994	Adz26994 Adeno-sss
42	2514.5	86.5	735	9	ADZ26999	Adz26999 Adeno-sss
43	2513.5	86.5	735	9	ADZ27035	Adz27035 Adeno-sss
44	2511	86.4	734	9	ADZ27037	Adz27037 Adeno-sss
45	2506.5	86.3	735	9	ADZ27006	Adz27006 Adeno-sss

ALIGNMENTS

RESULT 1
AAV71169
ID AAV71169 standard; protein; 534 AA.
XX
XX AAV71169;
AC
XX
XX
DT 08-SEP-2000 (first entry)
XX
XX

Adeno-associated virus serotype 1 capsid protein VP3.

Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy; vaccine; transgene; VP3.

Adeno-associated virus 1.

Adeno-associated virus 1.

Adeno-associated virus 1.

Adeno-associated virus 1.

Adeno-associated virus 1.

Adeno-associated virus 1.

Adeno-associated virus 1.

Adeno-associated virus 1.

Adeno-associated virus 1.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

CC medicament for the delivery of transgene to a host. The present sequence
CC is an AAV-1 cap protein VP3 which is useful in the production of
CC recombinant viral vector for gene delivery

xx Sequence 534 AA;

Query Match 100.0%; Score 2906; DB 3; Length 534;
Best Local Similarity 100.0%; Pred. No. 8e-236;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MASGGGAPMADNNEGADGVGNASGNWCHDSWTMLGDRVITTSRTTALPTYNHLYKQISS 60
DB 1 MASGGGAPMADNNEGADGVGNASGNWCHDSWTMLGDRVITTSRTTALPTYNHLYKQISS 60
QY 61 ASTGASNDNHYFGYSTPMGYPDENRFFCHFSRPMQRLINNMGFRPKRLNFKLNIQVK 120
DB 61 ASTGASNDNHYFGYSTPMGYPDENRFFCHFSRPMQRLINNMGFRPKRLNFKLNIQVK 120
QY 121 EVTTNDGVTITANNLTSTVOVFSDEYQLPYVLGSAHQCLPPPADVFMIPOGYLTIN 180
DB 121 EVTTNDGVTITANNLTSTVOVFSDEYQLPYVLGSAHQCLPPPADVFMIPOGYLTIN 180
QY 181 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYFEEVPHSSYAHSSQSLDRLMPLIDQY 240
DB 181 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYFEEVPHSSYAHSSQSLDRLMPLIDQY 240
QY 241 LYYLNRTONOGSGAQNKDLFSGSPAGMSVQPKMWLPQCYRQORVSKTNDNNSNFT 300
DB 241 LYYLNRTONOGSGAQNKDLFSGSPAGMSVQPKMWLPQCYRQORVSKTNDNNSNFT 300
QY 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 360
DB 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 360
QY 361 EEBIKATNPVATERFGTAVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 420
DB 361 EEBIKATNPVATERFGTAVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 420
QY 421 PHTDGHFHPSPLMGGFGKLNPPQILIKNTVPANPAPAEFSATKFASTIYOSTGOVSE 480
DB 421 PHTDGHFHPSPLMGGFGKLNPPQILIKNTVPANPAPAEFSATKFASTIYOSTGOVSE 480
QY 481 IEMELOKENSKRMPEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIRYLRPL 534
DB 481 IEMELOKENSKRMPEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIRYLRPL 534
```

RESULT 2
AAV71168
ID AAV71168 standard; protein; 599 AA.
XX

AC AAV71168;

DT 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 capsid protein VP2.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; VP2.

XX Adeno-associated virus 1.

OS

PN MO200028061-A2.

PD 18-MAY-2000.

PF 02-NOV-1999; 99WO-US025694.

PR 05-NOV-1998; 98US-0107114P.

XX (UYBE-) UNIV PENNSYLVANIA.
XX

PI Wilson JM, Xiao W;
XX WPI; 2000-376571/32.
DR N-PSDB; AAD00772, AAD00778.

PT Novel adeno-associated virus serotype 1 polynucleotide useful for
XX preparation of medicament for delivery of a transgene to a host.
PS Claim 7; Page 93-95; 108pp; English.

CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterized by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antisense delivery vectors. The AAV-1 does not induce the
CC formation of neutralising antibodies specific to any serotype of AAV
CC hence is useful for transforming host cells, and in the preparation of a
CC medicament for the delivery of transgene to a host. The present sequence
CC is an AAV-1 cap protein VP2 which is useful in the production of
CC recombinant viral vector for gene delivery

SQ Sequence 599 AA;

Query Match 100.0%; Score 2906; DB 3; Length 599;
Best Local Similarity 100.0%; Pred. No. 9.4e-236;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MASGGGAPMADNNEGADGVGNASGNWCHDSWTMLGDRVITTSRTTALPTYNHLYKQISS 60
DB 66 MASGGGAPMADNNEGADGVGNASGNWCHDSWTMLGDRVITTSRTTALPTYNHLYKQISS 125
QY 61 ASTGASNDNHYFGYSTPMGYPDENRFFCHFSRPMQRLINNMGFRPKRLNFKLNIQVK 120
DB 126 ASTGASNDNHYFGYSTPMGYPDENRFFCHFSRPMQRLINNMGFRPKRLNFKLNIQVK 185
QY 121 EVTTNDGVTITANNLTSTVOVFSDEYQLPYVLGSAHQCLPPPADVFMIPOGYLTIN 180
DB 186 EVTTNDGVTITANNLTSTVOVFSDEYQLPYVLGSAHQCLPPPADVFMIPOGYLTIN 245
QY 181 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYFEEVPHSSYAHSSQSLDRLMPLIDQY 240
DB 246 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYFEEVPHSSYAHSSQSLDRLMPLIDQY 305
QY 241 LYYLNRTONOGSGAQNKDLFSGSPAGMSVQPKMWLPQCYRQORVSKTNDNNSNFT 300
DB 306 LYYLNRTONOGSGAQNKDLFSGSPAGMSVQPKMWLPQCYRQORVSKTNDNNSNFT 365
QY 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 360
DB 366 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 425
QY 361 EEBIKATNPVATERFGTAVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 420
DB 426 EEBIKATNPVATERFGTAVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 485
QY 421 PHTDGHFHPSPLMGGFGKLNPPQILIKNTVPANPAPAEFSATKFASTIYOSTGOVSE 480
DB 486 PHTDGHFHPSPLMGGFGKLNPPQILIKNTVPANPAPAEFSATKFASTIYOSTGOVSE 545
QY 481 IEMELOKENSKRMPEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIRYLRPL 534
DB 546 IEMELOKENSKRMPEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIRYLRPL 599
```

RESULT 3
AAV71167
ID AAV71167 standard; protein; 736 AA.
XX
AC AAV71167;

XX 08-SEP-2000 (first entry)
 DT Adeno-associated virus serotype 1 capsid protein vp1.
 DE Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
 XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
 XX cap protein; recombinant viral vector; gene delivery; gene therapy;
 KM vaccine; transgene; VPI.
 XX Adeno-associated virus 1.
 OS WO200028061-A2.
 XX 18-MAY-2000.
 PD 02-NOV-1999; 99WO-US025694.
 XX 05-NOV-1998; 98US-0107114P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX Wilson JM, Xiao W;
 PI WPI; 2000-376571/32.
 DR N-PSDB; AAD00772, AAD00777.
 XX Novel adeno-associated virus serotype 1 polynucleotide useful for
 PT preparation of medicament for delivery of a transgene to a host.
 XX Claim 7, Page 87-90; 108pp; English.
 PS The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
 XX which is characterised by two inverted terminal repeats (ITR) and open
 CC reading frames for rep and capsid (cap) proteins. The rep reading frame
 CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
 CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
 CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
 CC regions, are useful in production of recombinant viral vectors for gene
 CC delivery. These vectors can be used as gene therapy vectors, vaccine
 CC vectors or antisense delivery vectors. The AAV-1 does not induce the
 CC formation of neutralising antibodies specific to any serotype of AAV
 CC hence is useful for transforming host cells, and in the preparation of a
 CC medicament for the delivery of transgene to a host. The present sequence
 CC is an AAV-1 cap protein VPI which is useful in the production of
 CC recombinant viral vector for gene delivery
 CC
 XX Sequence 736 AA;
 SQ
 Query Match 100.0%; Score 2906; DB 3; Length 736;
 Best Local Similarity 100.0%; Pred. No. 1.3e-235;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGAPADNNEGADGVGNASGNHCDSTWLGDRVITTTSTRTWALPTYNHLYKOISS 60
 DB 203 MASGGAPADNNEGADGVGNASGNHCDSTWLGDRVITTTSTRTWALPTYNHLYKOISS 262
 QY 61 ASTGASNDNHFGYSTPWFYDFENRPFCHSPRDWQRLINNNGFRKRLNFKLNFNIQVK 120
 DB 263 ASTGASNDNHFGYSTPWFYDFENRPFCHSPRDWQRLINNNGFRKRLNFKLNFNIQVK 322
 QY 121 EVTTNDGVTTIANNLSTVQVFSDEYQLPYVLSAHOGLPPPADVFMIPQGYLTIN 180
 DB 323 EVTTNDGVTTIANNLSTVQVFSDEYQLPYVLSAHOGLPPPADVFMIPQGYLTIN 382
 QY 181 NSGQAVRSGSPFCLFEPSPQMLRTGNNFTSYTFEEVPHSSVYHSSQSLDRMLNPLIDQY 240
 DB 383 NSGQAVRSGSPFCLFEPSPQMLRTGNNFTSYTFEEVPHSSVYHSSQSLDRMLNPLIDQY 442
 QY 241 LYYINRPNONGSGAKNDLLFSRGSFAGMSVQPKMILPGCYROQRVSKTKTDNNNSNFT 300
 DB 443 LYYINRPNONGSGAKNDLLFSRGSFAGMSVQPKMILPGCYROQRVSKTKTDNNNSNFT 502
 QY 301 WTGASKYTLNLRGSIINPGTAMASHKDEDEKFFPMGSGVMI FGKESAGASNTALDNVMTD 360

DB 503 WTGASKYTLNLRGSIINPGTAMASHKDEDEKFFPMGSGVMI FGKESAGASNTALDNVMTD 562
 QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMTQDRDVIYQGIWAKI 420
 DB 563 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMTQDRDVIYQGIWAKI 622
 QY 421 PHTDGHFHPSPLMGFGGLKNPPQQLIKNTVPAPPAEFSAFKFASFTQYSTGOVSVE 480
 DB 623 PHTDGHFHPSPLMGFGGLKNPPQQLIKNTVPAPPAEFSAFKFASFTQYSTGOVSVE 682
 QY 481 IEMELOKENSKRWNPEVOYTSNYAKSAAVDFVDNNGLYTEBRPIGTRYLTRPL 534
 DB 683 IEMELOKENSKRWNPEVOYTSNYAKSAAVDFVDNNGLYTEBRPIGTRYLTRPL 736

RESULT 4
 ABB80232
 ID ABB80232 standard; protein; 736 AA.
 XX ABB80232;
 XX 20-NOV-2003 (first entry)
 DT AAV1 vp1 protein.
 DE Adeno-associated virus; AAV; serotype 9; rep; cap; vp1; vp2; vp3;
 XX AAV splice variant; transgene.
 KM Adeno associated virus serotype 1.
 OS
 XX WO2003052052-A2.
 EN 26-JUN-2003.
 PD 12-NOV-2002; 2002WO-US033631.
 XX 17-DEC-2001; 2001US-0341150P.
 PR 05-JUN-2002; 2002US-0386132P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA Gao G, Wilson JM, Alvira M;
 PI WPI; 2003-523523/49.
 DR New isolated adeno-associated virus (AAV) comprising an AAV9 capsid,
 XX useful for preparing a medicament for delivering a transgene to a cell.
 PT Disclosure; Fig 2; 76pp; English.
 PS The sequences given in ABB80231-34 represent vp1 proteins derived from
 CC various adeno-associated virus (AAV) serotypes. These sequences were used
 CC in the scope of the invention for comparison with the cap protein derived
 CC from AAV serotype 9. The AAV capsid comprises three proteins vp1, vp2 and
 CC vp3, which are alternative splice variants. The AAV or the nucleic acid
 CC molecule is useful for preparing a medicament for delivering a transgene
 CC to a cell
 CC
 XX Sequence 736 AA;
 SQ
 Query Match 100.0%; Score 2906; DB 6; Length 736;
 Best Local Similarity 100.0%; Pred. No. 1.3e-235;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGAPADNNEGADGVGNASGNHCDSTWLGDRVITTTSTRTWALPTYNHLYKOISS 60
 DB 203 MASGGAPADNNEGADGVGNASGNHCDSTWLGDRVITTTSTRTWALPTYNHLYKOISS 262
 QY 61 ASTGASNDNHFGYSTPWFYDFENRPFCHSPRDWQRLINNNGFRKRLNFKLNFNIQVK 120
 DB 263 ASTGASNDNHFGYSTPWFYDFENRPFCHSPRDWQRLINNNGFRKRLNFKLNFNIQVK 322


```

Qy 121 EVTTNDGVTITANNLTSTVOVFSDESEYQLPYVLGSAHQCLPPPADVFMIPQGYTLTN 180
Db 323 EVTTNDGVTITANNLTSTVOVFSDESEYQLPYVLGSAHQCLPPPADVFMIPQGYTLTN 382
Qy 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEEVPFHSSYAHSSQSLDRLMPLIDQY 240
Db 383 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEEVPFHSSYAHSSQSLDRLMPLIDQY 442
Qy 241 LYYLNRTQNGSAGNKKDLFSGSPAGMSVQPKWMLPGPCYRQORVSKTIDNNNSNFT 300
Db 443 LYYLNRTQNGSAGNKKDLFSGSPAGMSVQPKWMLPGPCYRQORVSKTIDNNNSNFT 502
Qy 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGWIMFGKESAGASNTALDNVMTD 360
Db 503 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGWIMFGKESAGASNTALDNVMTD 562
Qy 361 EBEIKATNPVATERFGTVAANFQSSSTDPAIGDVHAMGALPGMWQODRDVYLQGPIMAKI 420
Db 563 EBEIKATNPVATERFGTVAANFQSSSTDPAIGDVHAMGALPGMWQODRDVYLQGPIMAKI 622
Qy 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAFSATKFSATITQYSTGQSVSE 480
Db 623 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAFSATKFSATITQYSTGQSVSE 682
Qy 481 IEMELQKENSKRMPNEVOYTSNYAKSANVDFTVNNGLYTEBRPIGTRYLTRPL 534
Db 683 IEMELQKENSKRMPNEVOYTSNYAKSANVDFTVNNGLYTEBRPIGTRYLTRPL 736

```

RESULT 5
ABR62762
ID ABR62762 standard; protein; 736 AA.

AC ABR62762;

DT 06-NOV-2003 (first entry)

DE Adeno associated virus 1 capsid protein vpl.

KM AAV, AAV1, capsid; vector; gene therapy; antisense therapy; vaccine.

OS Adeno associated virus.

PN WO2003052051-A2.

XX 26-JUN-2003.

XX 12-NOV-2002; 2002WO-US033630.

XX 17-DEC-2001; 2001US-0341151P.

PR 01-MAY-2002; 2002US-0377133P.

PR 05-JUN-2002; 2002US-0386122P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Gao G, Wilson JM, Alvira M;

PI MPI; 2003-523522/49.

XX New adeno-associated virus (AAV) comprising an AAV capsid having an amino

PT acid sequence of AAV8, useful for preparing a medicament for delivery of

PT a transgene to a cell and for treating cystic fibrosis or hemophilia B.

XX Disclosure; Fig 2A-C; 82pp; English.

CC The present sequence is the protein sequence of the vpl capsid protein of

CC adeno associated virus serotype 1 (AAV1). The invention provides the

CC nucleic acid and amino acid sequences of novel AAV8 and fragments of

CC these sequences. Each of these fragments may be used in a variety of

CC vector systems and host cells. Among the desirable fragments are the cap

CC proteins, including vpl, vp2, vp3 and hypervariable regions, the rep

CC proteins, including rep78, rep68, rep52 and rep40, and the sequences

CC encoding these proteins. The fragments may be used alone, in combination

CC with other AAV8 sequences or in combination with elements from other AAV

CC or non-AAV viral sequences in the production of recombinant AAV and for

CC use as antisense delivery vectors, gene therapy vectors or vaccine

CC vectors. A claimed molecule comprises a cap protein of a functional AAV

CC rep gene from a serotype selected from AAV1, AAV2, AAV3, AAV4, AAV5 and

CC AAV6

XX SQ Sequence 736 AA;

Query Match 100.0%; Score 2906; DB 7; Length 736;
Best Local Similarity 100.0%; Pred. No. 1,3e-235; Indels 0; Gaps 0;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MASGGA PMADNNGADGVGNASGNMHCDSITWLDRIYTTSTRTMALPTYNHLYKQISS 60
Db 203 MASGGA PMADNNGADGVGNASGNMHCDSITWLDRIYTTSTRTMALPTYNHLYKQISS 262
Qy 61 ASTGASNDNHFGYSTPMGYDFENRFCHFSRDMQRLINNMGFRPRKLNFKLFNIQVK 120
Db 263 ASTGASNDNHFGYSTPMGYDFENRFCHFSRDMQRLINNMGFRPRKLNFKLFNIQVK 322
Qy 121 EVTTNDGVTITANNLTSTVOVFSDESEYQLPYVLGSAHQCLPPPADVFMIPQGYTLTN 180
Db 323 EVTTNDGVTITANNLTSTVOVFSDESEYQLPYVLGSAHQCLPPPADVFMIPQGYTLTN 382
Qy 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEEVPFHSSYAHSSQSLDRLMPLIDQY 240
Db 383 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEEVPFHSSYAHSSQSLDRLMPLIDQY 442
Qy 241 LYYLNRTQNGSAGNKKDLFSGSPAGMSVQPKWMLPGPCYRQORVSKTIDNNNSNFT 300
Db 443 LYYLNRTQNGSAGNKKDLFSGSPAGMSVQPKWMLPGPCYRQORVSKTIDNNNSNFT 502
Qy 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGWIMFGKESAGASNTALDNVMTD 360
Db 503 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSGWIMFGKESAGASNTALDNVMTD 562
Qy 361 EBEIKATNPVATERFGTVAANFQSSSTDPAIGDVHAMGALPGMWQODRDVYLQGPIMAKI 420
Db 563 EBEIKATNPVATERFGTVAANFQSSSTDPAIGDVHAMGALPGMWQODRDVYLQGPIMAKI 622
Qy 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAFSATKFSATITQYSTGQSVSE 480
Db 623 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAFSATKFSATITQYSTGQSVSE 682
Qy 481 IEMELQKENSKRMPNEVOYTSNYAKSANVDFTVNNGLYTEBRPIGTRYLTRPL 534
Db 683 IEMELQKENSKRMPNEVOYTSNYAKSANVDFTVNNGLYTEBRPIGTRYLTRPL 736

```

RESULT 6
ADE76565
ID ADE76565 standard; protein; 736 AA.

AC ADE76565;

DT 29-JAN-2004 (first entry)

XX Adeno-associated virus (AAV) related protein, SEQ ID No 64.

XX adeno-associated virus; AAV; cytosolic; antiproliferative; antirheumatic;

XX antitachytic; neuroprotective; antidiabetic; antithyroid;

XX dermatological; antiinflammatory; gene therapy; vaccine;

XX hyperproliferative; cancer; psoriasis; autoimmune disease;

XX rheumatoid arthritis; multiple sclerosis; diabetes;

XX autoimmune thyroiditis; scleroderma; Crohn's disease.

XX Unidentified.

XX BP1310571-A2.

XX 14-MAY-2003.

SQ Sequence 736 AA;
 Query Match 100.0%; Score 2906; DB 9; Length 736;
 Best Local Similarity 100.0%; Pred. No. 1.3e-235;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTTSTRTWALPTYNHLYKQISS 60
 DB 203 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTTSTRTWALPTYNHLYKQISS 262
 QY 61 ASTGASNDNHFGYSTPMGYFDENRPFCHSPRDMORLNNMGFRPRKLFKLFNIOVK 120
 DB 263 ASTGASNDNHFGYSTPMGYFDENRPFCHSPRDMORLNNMGFRPRKLFKLFNIOVK 322
 QY 121 EYTTNDGVTITANNLITSTVQVFSDEBYQLPYVLGSAHQGCLPPPADVFMIPOYGYLTLN 180
 DB 323 EYTTNDGVTITANNLITSTVQVFSDEBYQLPYVLGSAHQGCLPPPADVFMIPOYGYLTLN 382
 QY 181 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYASHSOSLDRLMNPIDQY 240
 DB 363 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYASHSOSLDRLMNPIDQY 442
 QY 241 LYYLNRTONOGSGAQNKDILFSGSPAGMSVOPKMWLPGPCYRQORVSKTDTNNNSFT 300
 DB 443 LYYLNRTONOGSGAQNKDILFSGSPAGMSVOPKMWLPGPCYRQORVSKTDTNNNSFT 502
 QY 301 WTGASKYVNLNGRESIINPGTAMASHKDEDEKFFPMGVMIPGKESAGASNTALDNVMTTD 360
 DB 503 WTGASKYVNLNGRESIINPGTAMASHKDEDEKFFPMGVMIPGKESAGASNTALDNVMTTD 562
 QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 420
 DB 563 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 622
 QY 421 PHTDGHFHPSPLMGFGGLKNPPPOILLKNTVPANPPAEFSATKASFTTQYSTQGVSE 480
 DB 623 PHTDGHFHPSPLMGFGGLKNPPPOILLKNTVPANPPAEFSATKASFTTQYSTQGVSE 682
 QY 481 IEMELOKENSKRMNPEVOYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 534
 DB 683 IEMELOKENSKRMNPEVOYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 736
 RESULT 8
 AD227069
 ID AD227069 standard; protein; 736 AA.
 AC AD227069;
 XX
 DT 30-JUN-2005 (first entry)
 XX
 DE Adeno-associated virus protein SEQ ID NO 219.
 XX
 XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KW bacterial infection; cancer; ulcerative colitis; antihemuratic;
 KW antiarthritis; neuroprotective; antiinflammatory; antidiabetic;
 KW antipneumatic; vasotropic; gastrointestinal gen.; hemostatic; anti-HIV;
 KW virucide; antibacterial; cytostatic; anticancer; dermatological.
 OS
 XX Adeno-associated virus.
 PN
 XX MO2005033321-A2.
 PD 14-APR-2005.
 XX
 XX 30-SEP-2004; 2004MO-US028817.
 XX
 XX 30-SEP-2003; 2003US-0508226P.
 PR 29-APR-2004; 2004US-0566546P.
 XX
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA
 XX

PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
 XX
 DR WPI; 2005-285437/29.
 XX
 PT New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.
 XX
 PS Disclosure; SEQ ID NO 219; 569bp; English.
 XX
 CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC scleroderma, diabetes, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.
 SQ
 SQ Sequence 736 AA;
 Query Match 100.0%; Score 2906; DB 9; Length 736;
 Best Local Similarity 100.0%; Pred. No. 1.3e-235;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTTSTRTWALPTYNHLYKQISS 60
 DB 203 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTTSTRTWALPTYNHLYKQISS 262
 QY 61 ASTGASNDNHFGYSTPMGYFDENRPFCHSPRDMORLNNMGFRPRKLFKLFNIOVK 120
 DB 263 ASTGASNDNHFGYSTPMGYFDENRPFCHSPRDMORLNNMGFRPRKLFKLFNIOVK 322
 QY 121 EYTTNDGVTITANNLITSTVQVFSDEBYQLPYVLGSAHQGCLPPPADVFMIPOYGYLTLN 180
 DB 323 EYTTNDGVTITANNLITSTVQVFSDEBYQLPYVLGSAHQGCLPPPADVFMIPOYGYLTLN 382
 QY 181 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYASHSOSLDRLMNPIDQY 240
 DB 383 NGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVPHSSYASHSOSLDRLMNPIDQY 442
 QY 241 LYYLNRTONOGSGAQNKDILFSGSPAGMSVOPKMWLPGPCYRQORVSKTDTNNNSFT 300
 DB 443 LYYLNRTONOGSGAQNKDILFSGSPAGMSVOPKMWLPGPCYRQORVSKTDTNNNSFT 502
 QY 301 WTGASKYVNLNGRESIINPGTAMASHKDEDEKFFPMGVMIPGKESAGASNTALDNVMTTD 360
 DB 503 WTGASKYVNLNGRESIINPGTAMASHKDEDEKFFPMGVMIPGKESAGASNTALDNVMTTD 562
 QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 420
 DB 563 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVTYLOGPIMAKI 622
 QY 421 PHTDGHFHPSPLMGFGGLKNPPPOILLKNTVPANPPAEFSATKASFTTQYSTQGVSE 480
 DB 623 PHTDGHFHPSPLMGFGGLKNPPPOILLKNTVPANPPAEFSATKASFTTQYSTQGVSE 682
 QY 481 IEMELOKENSKRMNPEVOYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 534
 DB 683 IEMELOKENSKRMNPEVOYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 736
 RESULT 9
 AD227012
 ID AD227012 standard; protein; 736 AA.
 AC AD227012;
 XX
 DT 30-JUN-2005 (first entry)
 XX

XX Adeno-associated virus protein SEQ ID NO 162.
 DE rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 XX bacterial infection; cancer; ulcerative colitis; antirheumatic;
 XX antiarthritic; neutropenic; antidiabetic;
 XX antiparasitic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 XX virucide; antibacterial; cytostatic; antitumor; dermatological.
 OS Adeno-associated virus.
 XX WO200503321-A2.
 XX 14-APR-2005.
 XX 30-SEP-2004; 2004WO-US028817.
 XX 30-SEP-2003; 2003US-0508226P.
 XX 29-APR-2004; 2004US-0566546P.
 XX (UYPE-) UNIT PENNSYLVANIA.
 XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
 PI WPI; 2005-285437/29.
 XX New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.
 XX Disclosure; SEQ ID NO 162; 569pp; English.
 XX The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.
 XX Sequence 736 AA;
 SQ
 Query Match 100.0%; Score 2906; DB 9; Length 736;
 Best Local Similarity 100.0%; Pred. No. 1.3e-235;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGAPMADNNEGADGVGNAGNHCDSITWIGDRITTTSTTWMALPTNNHLYQISS 60
 DB 203 MASGGAPMADNNEGADGVGNAGNHCDSITWIGDRITTTSTTWMALPTNNHLYQISS 262
 QY 61 ASGASNDNHYFGYSTPMGFDFRFRCHSPRDMQRLINNNNGFRKRLNFKLNIOVK 120
 DB 263 ASGASNDNHYFGYSTPMGFDFRFRCHSPRDMQRLINNNNGFRKRLNFKLNIOVK 322
 QY 121 EYTTNDGVTTIANNLSTVQVPSDSEYQLPYVIGSAHQGLPPPADVPMIPOGYVLTIN 180
 DB 323 EYTTNDGVTTIANNLSTVQVPSDSEYQLPYVIGSAHQGLPPPADVPMIPOGYVLTIN 382
 QY 181 NGGQAVGRSSPYCLIEFPSCMLRGNFTSYTTEEPFSSVAHSSLSRLNMLDYOY 240
 DB 383 NGGQAVGRSSPYCLIEFPSCMLRGNFTSYTTEEPFSSVAHSSLSRLNMLDYOY 442
 QY 241 LYYNTRTONSGSQQNDLFLSRGSPAGMSVQPKMPLPGPCYRQOVSXTKTNDNNSNFT 300
 DB 443 LYYNTRTONSGSQQNDLFLSRGSPAGMSVQPKMPLPGPCYRQOVSXTKTNDNNSNFT 502
 QY 301 WTGASKYLNLRGRESIINPGTAMASHKDEDFPFMSGVMIFFKESAGASNTALDNVMTD 360

DB 503 WTGASKYLNLRGRESIINPGTAMASHKDEDFPFMSGVMIFFKESAGASNTALDNVMTD 562
 QY 361 EEEIKATNPVATERFGVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVIYQGPIMAKI 420
 DB 563 EEEIKATNPVATERFGVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVIYQGPIMAKI 622
 QY 421 PRTDGHFHPSPIMGSPGLKNPPQILINTPVPANPAPAEFSATKFSFIYOSTGOVSVE 480
 DB 623 PRTDGHFHPSPIMGSPGLKNPPQILINTPVPANPAPAEFSATKFSFIYOSTGOVSVE 682
 QY 481 IEMELQKNSKRWNPDEVQYTSNYAKSANVDFVDNNGLYTEBRPIGTRYLTRPL 534
 DB 683 IEMELQKNSKRWNPDEVQYTSNYAKSANVDFVDNNGLYTEBRPIGTRYLTRPL 736
 RESULT 10
 ADV67506
 ID ADV67506 standard; protein; 735 AA.
 XX ADV67506;
 XX 10-MAR-2005 (first entry)
 XX Amino acid sequence of the capsid protein of AAV serotype 1.
 XX antiarteriosclerotic; antidiabetic; gene therapy; cholesterol;
 XX apolipoprotein E; apoB; apolipoprotein A; apoA; atherosclerosis;
 XX lipoprotein defect; capsid protein.
 OS Adeno-associated virus.
 XX WO2004108922-A2.
 XX 16-DEC-2004.
 XX 23-APR-2004; 2004WO-US010965.
 XX 25-APR-2003; 2003US-0465293P.
 XX (UYPE-) UNIT PENNSYLVANIA.
 XX Rader DJ, Wilson JM;
 PI WPI; 2005-031700/03.
 XX Lowering total cholesterol levels and treating atherosclerosis in a
 PT subject comprises delivering a recombinant adeno-associated virus (AAV)
 PT comprising an AAV serotype capsid protein or a gene encoding human
 PT apolipoprotein E (apoB) or apoA.
 PS Disclosure; SEQ ID NO 4; 69pp; English.
 XX The specification describes a method for lowering total cholesterol
 CC levels in a subject. The method comprises delivering to the subject a
 CC recombinant adeno-associated virus (AAV) comprising a gene encoding a
 CC human apolipoprotein E (apoB) or apoA under the control of a regulatory
 CC control sequence which directs expression of the gene. The recombinant
 CC AAV also comprises a capsid protein selected from an AAV serotype (e.g.,
 CC AAV7 or AAV8) which preferentially expresses high levels of transgene in
 CC live. A therapeutically effective amount of apoB or apoA expression is
 CC obtained upon delivery of low dose of AAV. The method of the invention is
 CC useful for lowering total cholesterol levels in a subject, e.g., for
 CC treating atherosclerosis, and for correcting defects in lipoprotein. The
 CC present sequence represents a capsid protein of AAV serotype 1 (AAV1),
 CC which may be used in recombinant AAV vectors of the invention.
 XX Sequence 735 AA;
 SQ
 Query Match 99.9%; Score 2902; DB 9; Length 735;
 Best Local Similarity 100.0%; Pred. No. 2.8e-235;
 Matches 533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGAGPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 60
Db 203 MASGGAGPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPMGYDFDNRFHCHFSFPRDMQRLINNMGFRPRRLNFKLFNIQVK 120
Db 263 ASTGASNDNHFGYSTPMGYDFDNRFHCHFSFPRDMQRLINNMGFRPRRLNFKLFNIQVK 322
QY 121 EVTTNDGVTITANNLITSTVQVFSDEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTLN 180
Db 323 EVTTNDGVTITANNLITSTVQVFSDEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTLN 382
QY 181 NGSQAVGRSSFFCYCLEYFSPSQMLRTGNNTFFSYTFEEVPHSSVYAHOSLDRIMPDLIOY 240
Db 363 NGSQAVGRSSFFCYCLEYFSPSQMLRTGNNTFFSYTFEEVPHSSVYAHOSLDRIMPDLIOY 442
QY 241 LYYLNRTONOGSSAONKDLFSRGS PAKMSVOPKMWLPQPCYRQORVSKTTDNNNSNFT 300
Db 443 LYYLNRTONOGSSAONKDLFSRGS PAKMSVOPKMWLPQPCYRQORVSKTTDNNNSNFT 502
QY 301 WTGASKYVLNGRRESITINPGTAMASHKDEDEKFFPMGVMIFGKSASAGSNTALDNVMTTD 360
Db 503 WTGASKYVLNGRRESITINPGTAMASHKDEDEKFFPMGVMIFGKSASAGSNTALDNVMTTD 562
QY 563 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQODRDVYLQGIWAKI 420
Db 622 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQODRDVYLQGIWAKI 622
QY 421 PHTDGHFHPSPLMGFGKLPNPOLLIKNTVPANPAPAFSATKASFTITQYSTQGVSE 480
Db 623 PHTDGHFHPSPLMGFGKLPNPOLLIKNTVPANPAPAFSATKASFTITQYSTQGVSE 682
QY 481 IEMELQKENSKRMPNEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIVYLRPL 533
Db 683 IEMELQKENSKRMPNEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIVYLRPL 735

RESULT 11
AD227008
ID AD227008 standard; protein, 736 AA.
XX
AC AD227008;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 158.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; anti-inflammatory;
KW antiarthritis; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotrophic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS Adeno-associated virus.
XX
PN MO200503321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
XX
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
DR WPI, 2005-285437/29.
XX
PT New adeno-associated virus (AAV) clade comprising at least three AAV
members, useful for preventing and/or treating arthritis, multiple

PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
PS Claim 19; SEQ ID NO 158; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
SQ Sequence 736 AA;
XX
Query Match 99.5%; Score 2892; DB 9; Length 736;
Best Local Similarity 99.6%; Pred. No. 1.9e-234;
Matches 532; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MASGGAGPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 60
Db 203 MASGGAGPMADNNEGADGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPMGYDFDNRFHCHFSFPRDMQRLINNMGFRPRRLNFKLFNIQVK 120
Db 263 ASTGASNDNHFGYSTPMGYDFDNRFHCHFSFPRDMQRLINNMGFRPRRLNFKLFNIQVK 322
QY 121 EVTTNDGVTITANNLITSTVQVFSDEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTLN 180
Db 323 EVTTNDGVTITANNLITSTVQVFSDEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTLN 382
QY 181 NGSQAVGRSSFFCYCLEYFSPSQMLRTGNNTFFSYTFEEVPHSSVYAHOSLDRIMPDLIOY 240
Db 363 NGSQAVGRSSFFCYCLEYFSPSQMLRTGNNTFFSYTFEEVPHSSVYAHOSLDRIMPDLIOY 442
QY 241 LYYLNRTONOGSSAONKDLFSRGS PAKMSVOPKMWLPQPCYRQORVSKTTDNNNSNFT 300
Db 443 LYYLNRTONOGSSAONKDLFSRGS PAKMSVOPKMWLPQPCYRQORVSKTTDNNNSNFT 502
QY 503 WTGASKYVLNGRRESITINPGTAMASHKDEDEKFFPMGVMIFGKSASAGSNTALDNVMTTD 360
Db 622 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQODRDVYLQGIWAKI 420
QY 421 PHTDGHFHPSPLMGFGKLPNPOLLIKNTVPANPAPAFSATKASFTITQYSTQGVSE 480
Db 623 PHTDGHFHPSPLMGFGKLPNPOLLIKNTVPANPAPAFSATKASFTITQYSTQGVSE 682
QY 481 IEMELQKENSKRMPNEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIVYLRPL 534
Db 683 IEMELQKENSKRMPNEVOYTSNYAKSANDFTVDNNGLYTEBRPIGTIVYLRPL 736

RESULT 12
AD227086
ID AD227086 standard; protein, 737 AA.
XX
AC AD227086;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 236.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antiinflammatory;

KM antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
 KM antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 KM virucide; antibacterial; cyostatic; antitumor; dermatological.
 XX
 OS Adeno-associated virus.
 XX
 FN WO2005033321-A2.
 XX
 PD 14-APR-2005.
 XX
 PF 30-SEP-2004; 2004WO-US028817.
 XX
 PR 30-SEP-2003; 2003US-0508226P.
 PR 29-APR-2004; 2004US-0566546P.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 XX
 PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
 XX
 DR WPI; 2005-285437/29.
 XX
 XX New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.
 XX
 PS Claim 19; SEQ ID NO 236; 569bp; English.
 XX
 CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.
 CC
 XX
 SQ Sequence 737 AA;
 Query Match 99.4%; Score 2892; DB 9; Length 737;
 Best Local Similarity 99.4%; Pred. No. 1.9e-234;
 Matches 531; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 60
 DB 204 MASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 263
 QY 61 ASAGASNDNHYFGYSTPMGYFDFNRFCHSPSPDMQRLINNNNGFRKRLNPLFNIOVK 120
 DB 264 ASAGASNDNHYFGYSTPMGYFDFNRFCHSPDMQRLINNNNGFRKRLNPLFNIOVK 323
 QY 121 EVTTNDGVTTIANNLSTGVVFSDESEYDLPVYLSAHOGLPPPADVEMIPQYGYLTN 180
 DB 324 EVTTNDGVTTIANNLSTGVVFSDESEYDLPVYLSAHOGLPPPADVEMIPQYGYLTN 383
 QY 181 NSQAVGRSSFFCLIEFTPSQMLRTGNFTSYTFEEVPLHSSVAHSQSLDRLNPLIDQY 240
 DB 384 NSQAVGRSSFFCLIEFTPSQMLRTGNFTSYTFEEVPLHSSVAHSQSLDRLNPLIDQY 443
 QY 241 LYYLNRTONOSGSAONKDLFSRGSFAGMSVOPKMLPGCYRQORVSKTTDNNSNFT 300
 DB 444 LYYLNRTONOSGSAONKDLFSRGSFAGMSVOPKMLPGCYRQORVSKTTDNNSNFT 503
 QY 301 WTGASKYNLNGRESIINPGTAMASHKODEKFFPMGSMVIFGKESAGASTALDNWITD 360
 DB 504 WTGASKYNLNGRESIINPGTAMASHKODEKFFPMGSMVIFGKESAGASTALDNWITD 563
 QY 361 EEEFKANNPVATREFGVAVNFOSSSTDPATGDVHAMGALPGMWODRDVYLGSPITAKI 420
 DB 564 EEEFKANNPVATREFGVAVNFOSSSTDPATGDVHAMGALPGMWODRDVYLGSPITAKI 623

QY 421 PHTDGHFSPPLMGSGFLKNNPPOLIKNTVPANPPAEFSATKFPASFILOYSGOVSE 480
 DB 624 PHTDGHFSPPLMGSGFLKNNPPOLIKNTVPANPPAEFSATKFPASFILOYSGOVSE 683
 QY 481 IEMELOKENSKRMNEVOYTSNYAKSANVDFVDDNNGLYTEPRPIGTRYLTPPL 534
 DB 684 IEMELOKENSKRMNEVOYTSNYAKSANVDFVDDNNGLYTEPRPIGTRYLTPPL 737
 RESULT 13
 AD227074
 ID AD227074 strand; protein; 736 AA.
 XX
 AC AD227074;
 XX
 DT 30-JUN-2005 (first entry)
 XX
 DE Adeno-associated virus protein SEQ ID NO 224.
 XX
 KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
 KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
 KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 KW virucide; antibacterial; cyostatic; antitumor; dermatological.
 XX
 OS Adeno-associated virus.
 XX
 PN WO2005033321-A2.
 XX
 PD 14-APR-2005.
 XX
 PF 30-SEP-2004; 2004WO-US028817.
 XX
 PR 30-SEP-2003; 2003US-0508226P.
 PR 29-APR-2004; 2004US-0566546P.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 XX
 PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
 XX
 DR WPI; 2005-285437/29.
 XX
 XX New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.
 XX
 PS Claim 19; SEQ ID NO 224; 569bp; English.
 XX
 CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.
 CC
 XX
 SQ Sequence 736 AA;
 Query Match 99.4%; Score 2888; DB 9; Length 736;
 Best Local Similarity 99.4%; Pred. No. 4.2e-234;
 Matches 531; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 60
 DB 203 MASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 262

Job time : 97.4286 secs

```
XX 05-DEC-2000.
PD 11-JUN-1997; 97US-00873168.
XX 11-JUN-1997; 97US-00873168.
XX 11-JUN-1997; 97US-00873168.
XX (UNITW ) UNITV WASHINGTON.
XX PA
XX PI
XX Ruseell DW, Rutledge EA;
XX WPI; 2001-060164/07.
XX
XX Adeno-associated virus serotype 6 and viral vector derived from it for
XX PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
XX PT syndrome, sickle cell anemia, thalassemia and diabetes.
XX
XX PS Disclosure; Fig 2; 50pp; English.
XX
XX CC The present invention relates to adeno-associated virus serotypes. The
XX CC present sequence is capsid protein VP1 of one such serotype (AAV6). AAV6
XX CC can be used to construct AAV viral vectors for use in gene therapy for a
XX CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
XX CC sickle cell anaemia, thalassemia, blood clotting disorders and diabetes.
XX CC The AAV viral vectors have increased transduction efficiency of a
XX CC particular host cell as the AAV virion containing the AAV vector genome
XX CC can be modified to express a capsid protein of an AAV serotype that
XX CC transduces the selected host cell
XX
XX SQ Sequence 736 AA;
XX
XX Query Match 99.2%; Score 2884; DB 4; Length 736;
XX Best Local Similarity 99.1%; Pred. No. 9.1e-234;
XX Matches 529; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 MASGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
XX DB 203 MASGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 262
XX
XX QY 61 ASTGASNDNHYPFGYSTPMWGFDFNRFPHCHSPRDWQRLINNNMGFRPKRLNFKLFNIQYK 120
XX DB 263 ASTGASNDNHYPFGYSTPMWGFDFNRFPHCHSPRDWQRLINNNMGFRPKRLNFKLFNIQYK 322
XX
XX QY 121 EYTTNDGVTTIANNLTSTVOVFSDEXOLPYVLGSAHQCLPPPADVFMIPOYGYLTIN 180
XX DB 323 EYTTNDGVTTIANNLTSTVOVFSDEXOLPYVLGSAHQCLPPPADVFMIPOYGYLTIN 382
XX
XX QY 181 NSQAVGRSSFCLEYFPSCMLRTGNNFTFSYTFEEVPHHSYAHQSOLDRLNPLIDQY 240
XX DB 383 NSQAVGRSSFCLEYFPSCMLRTGNNFTFSYTFEEVPHHSYAHQSOLDRLNPLIDQY 442
XX
XX QY 241 LYYLNTQONQSGAONKDLLEFSRGSPPAGMSVOPKNNLPGPCYRQORVSKTKTDNNNSNFT 300
XX DB 443 LYYLNTQONQSGAONKDLLEFSRGSPPAGMSVOPKNNLPGPCYRQORVSKTKTDNNNSNFT 502
XX
XX QY 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCVMI FGKESAGASNTALDNWMTD 360
XX DB 503 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCVMI FGKESAGASNTALDNWMTD 562
XX
XX QY 361 EEEIKATNPVATERFGTVAVNFOSSSTDPAITGDVHAMGALPGMWQDRDYLLOGPIWAKI 420
XX DB 563 EEEIKATNPVATERFGTVAVNFOSSSTDPAITGDVHAMGALPGMWQDRDYLLOGPIWAKI 622
XX
XX QY 421 PHTDGHFSPPLMGFGFLKNPPOIILKNTPVANPPAEPASATKFAFITQYSTGQVSV 480
XX DB 623 PHTDGHFSPPLMGFGFLKNPPOIILKNTPVANPPAEPASATKFAFITQYSTGQVSV 682
XX
XX QY 481 IEWELQKENSKRNNPEVOYTSNVAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
XX DB 683 IEWELQKENSKRNNPEVOYTSNVAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:29:19 / Search time 18.8571 Seconds
(without alignments)
2724.684 Million cell updates/sec

Title: US-10-696-282-17

Sequence: 1 MASGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2177.5	74.9	504	1 VCPV3A	coat protein - ade
2	1676	57.7	732	2 S52210	coat protein VP1 -
3	517	17.8	673	1 VCPVB5	coat protein VP1 -
4	487.5	16.8	781	1 VCPV19	coat protein VP1 -
5	234.5	8.1	723	1 VCPVPP	coat protein VP1 -
6	232.5	8.0	729	1 A60006	coat protein VP1 -
7	232.5	8.0	729	1 VCPVNA	coat protein VP1 -
8	229	7.9	587	1 B44276	coat protein VP1 -
9	226	7.8	722	1 VCPV2	coat protein VP1 -
10	197	6.8	716	1 VCPV2M	coat protein VP1 -
11	196.5	6.8	718	1 VCPVIM	coat protein VP1 -
12	194	6.7	722	1 VCPVME	coat protein VP1 -
13	194	6.7	727	1 VCPVIF	coat protein VP1 -
14	189	6.5	584	2 S49594	capsid protein VP2
15	189	6.5	727	1 VCPVFP	coat protein VP1 -
16	188	6.5	737	1 VCPVCD	coat protein VP1 -
17	184	6.3	748	1 VCPVCP	coat protein VP1 -
18	177.5	6.1	702	1 VCPVAP	coat protein VP1 -
19	175	6.0	722	1 VCPVCN	coat protein VP1 -
20	123.5	4.2	680	2 AB0124	probable TonB-depe
21	122.5	4.2	880	1 SYBSYS	valine-RNA ligase
22	121	4.2	735	2 F56660	protein F2K1.10
23	118	4.1	739	2 T52289	probable transketo
24	115.5	4.0	345	1 B97024	probable phosphoe
25	113.5	3.9	2817	2 B97033	uncharacterized pr
26	109	3.8	418	2 T35753	probable periplasm
27	109	3.8	1072	2 A86827	hypothetical prote
28	109	3.8	1186	2 T18210	delta endotoxin -
29	108.5	3.7	1711	2 AB1283	peptidoglycan link

30	108	3.7	655	1 ALK6G	cyclomaldextrin
31	108	3.7	772	1 A44052	outer layer protei
32	108	3.7	826	2 AD1683	penicillin-binding
33	107.5	3.7	413	2 T23098	hypothetical prote
34	107.5	3.7	648	2 S50856	whn protein - rat
35	107	3.7	777	2 AE0462	probable exported
36	107	3.7	1338	2 T30565	MAP kinase kinase
37	106.5	3.7	403	2 B87513	acyl-CoA dehydroge
38	106	3.6	1271	2 D64237	hypothetical prote
39	105.5	3.6	2500	2 G71609	hypothetical prote
40	105	3.6	1394	2 A29637	position-specific
41	104.5	3.6	559	2 S33724	transcription fact
42	104.5	3.6	717	2 H85171	DEF (CIA1) protein
43	104	3.6	395	2 S50986	MAP1 protein - yea
44	104	3.6	715	2 H90977	hypothetical prote
45	104	3.6	719	2 F85824	hypothetical prote

ALIGNMENTS

RESULT 1

VCPV3A

coat protein - adeno-associated virus type 2

C:Species: adeno-associated virus type 2
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004

C:Accession: A03698

R:Srivastava, A.; Lusby, E.W.; Berns, K.I.

J. Virol. 45, 555-564, 1983

A:Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.

A:Reference number: A03694; MUID:83164299; PMID:6300419

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-504 <SRI>

A:Cross-references: UNIPROT:P03135; UNIPARC:UPI0000127C8E; EMBL:J01901; NID:G209616; PDB

C:Superfamily: adeno-associated virus coat protein

C:Keywords: coat protein

Query Match 74.9%; Score 2177.5; DB 1; Length 504;

Best Local Similarity 82.5%; Pred. No. 1.4e-152;

Matches 392; Conservative 33; Mismatches 47; Indels 3; Gaps 2;

QY	1	MASGGAPMADNNEGADGVGNASGWHCDSTWLGDRVITTSRTMALPTYNHLYKQISS	60
DB	1	MATGSGAPMADNNEGADGVGNSSGWHCDSTWMDGRVITTSRTMALPTYNHLYKQISS	60
QY	61	ASTGASNDNHYRGYSTPMPGYPDPNRFCHFSFRDMORLINNMWGRPRKLNFKLFNIQVK	120
DB	61	QS-GASNDNHYRGYSTPMPGYPDPNRFCHFSFRDMORLINNMWGRPRKLNFKLNFIQVK	119
QY	121	EYTTNDGVTTIANNTSTVQVFSDEYOLPYVLGSAHQCLPPPADVEMIDQYGLTLN	180
DB	120	EYTTNDGVTTIANNTSTVQVFTDSEYOLPYVLGSAHQCLPPPADVEMIDQYGLTLN	179
QY	181	NSQAVGRSSFYCLEYFPGOMLRTGNFTFSTFEEVPHSSYASQSIDRLMPLIDY	240
DB	180	NSQAVGRSSFYCLEYFPGOMLRTGNFTFSTFEEVPHSSYASQSIDRLMPLIDY	239
QY	241	LYLINRTQNSGSAQKDLFSRGSFAGMSVDPKQMLPGPCRCROORVSKTKTDNNNSFT	300
DB	240	LYLINRTNTPSGTITTSQSRLOFSQAGASDIRDSRWMLPGPCRCROORVSKTKTDNNNSFT	299
QY	301	WTGASKYKLNNGRESIINPGTAMASHKDDKFFPMGSMVIFPKESAGSNTALDVMYTD	360
DB	300	WTGATKYLHNGRDSLVP--AMASHKDDKFFPQSGVLIFPKQSEKKNVNIIEKVMYTD	357
QY	361	EEBIAKTNPVATERFCTAVNFQSSSTDPATGDFHAMGALPGMVMODRVRVYQGPYIAKI	420
DB	358	EEBIEGTTNPVATEQGVSTNLQRGNRQATADVTVOGLPGMVMODRVRVYQGPYIAKI	417
QY	421	PHTDGHFHSPLMGSGGLKNPPPOLIKNTPPVAPNPAEFSATKFASTFYOSTG	475
DB	421	PHTDGHFHSPLMGSGGLKNPPPOLIKNTPPVAPNPAEFSATKFASTFYOSTG	475

Db 418 PHTDGHFHSPLMGFGFKHPPLQILIKNTVPANPSTTFSNAKPFSTIOTYSTG 472

RESULT 2

S52210

coat protein VP1 - muscovy duck parvovirus

N/Alternate names: VP1 protein

C/Species: muscovy duck parvovirus

C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C/Accession: S52210

R/Zadori, Z.; Eidel, J.; Nagy, J.; Kisari, J.

Submitted to the EMBL Data Library, September 1993

A/Reference number: S52209

A/Accession: S52210

A/Molecule type: DNA

A/Residues: 1-732 <ZAD>

A/Cross-References: UNIPROT:O83289; UNIPARC:UPI0000066CSD; EMBL:X75093; NID:g609091; PDB

A/Experimental source: strain FM

C/Genetics:

A/Gene: VP1

C/Superfamily: parvovirus coat protein

C/Keywords: coat protein

Query Match 57.7%; Score 1676; DB 2; Length 732;

Best Local Similarity 56.9%; Pred. No. 2,1e-115;

Matches 308; Conservative 77; Mismatches 142; Indels 14; Gaps 6;

QY 1 MASGGAPMADNNEGADGVGNASGMHCDSTWLDGRTVITSTRWALPTYNHLYKOISS 60

Db 199 MAEGSGAGAGDAGADGVGNASGMHCDSTWLDGRTVITSTRWALPTYNHLYKOISS 258

QY 61 ASTGASNDNHVFGYSTPWCYFPDNRFHCHFSRPRDQRLINNMGFRPKLNFNIQVK 120

Db 259 GINPDSN-TQVAGYSTPWCYFPDNRFHCHFSRPRDQRLINNMGFRPKLNFNIQVK 317

QY 121 EYTTNDGVTTIANNTLSTVQVFSDBEYQLPYVLSAHQGLCPPADVFMIPOYGYLTN 180

Db 318 EYTTDQGTIANNTLSTIQTFTDNEHQLPYVLSAHQGLCPPADVFMIPOYGYCTMH 377

QY 181 ---NSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSVYASQSLDRL 237

Db 378 TNQSGARFNDRAAFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSVYASQSLDRL 437

QY 238 DQYLYLNRTQNSGSAQNKDLLFSRGPAGMSVQPKMLPGPCYRQORVSKTK--TDNN 295

Db 438 DQYLYNFSFV-NGGNNAQ-----FKKAVYGAGFGAGGRMLPSPKLLDQRYRAYSGCTNY 491

QY 236 NSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCVMTFGKE--SAGASNTAL 353

Db 492 ANMSIMSKGNKVFLLKDRYLQPGPVALTHTEDQASSVPAQNIIGIAKDPYRSGSTLAGI 551

QY 354 DNVMITDEEIKATNPVALTERFGYAVVNPQSSSTPACDVAAMGALGMQMODDVLQ 413

Db 552 SGMITVDEEIAPTLVGWRPYGLTVYNNQNTTAPVNALEVLDALEPMWQNDIYLO 611

QY 414 GPIMAKIPTDGHFHSPLMGFGFKHPPLQILIKNTVPANPAPAEFSATKPFASITQYS 473

Db 612 GPIMAKIPTDGHFHSPLMGFGFKHPPLQILIKNTVPANPAPAEFSATKPFASITQYS 671

QY 474 TQGVSEIEMELOKENSKRNPDEVQYTSNVAKSANDFTVDNNGLYTEPRPIGTRYLRP 533

Db 672 TQGVSEIEMELOKENSKRNPDEVQYTSNVAKSANDFTVDNNGLYTEPRPIGTRYLRP 731

QY 534 L 534

Db 732 L 732

RESULT 3

VCPVBS

coat protein VP1 - bovine parvovirus

N/Contains: coat protein VP2

C/Species: bovine parvovirus

C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004

C/Accession: A26104

R/Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.

J. Virol. 60, 1085-1097, 1986

A/Title: Complete nucleotide sequence and genome organization of bovine parvovirus.

A/Reference number: A26104; MUID:87061184; PMID:3783814

A/Accession: A26104

A/Molecule type: DNA

A/Residues: 1-673 <CHB>

A/Cross-References: UNIPROT:P07297; UNIPARC:UPI0000127D77; EMBL:M14363; NID:G333454; PDB

C/Superfamily: parvovirus coat protein

C/Keywords: coat protein

F/138-673/Product: coat protein VP2 #status predicted <VP2>

Query Match 17.8%; Score 517; DB 1; Length 673;

Best Local Similarity 25.1%; Pred. No. 4.4e-30;

Matches 143; Conservative 85; Mismatches 208; Indels 134; Gaps 17;

QY 5 GGAPMADNNEGADGVGNASGMHCDSTWLDGRTVITSTRWALPTYNHLYKOISSASTG 64

Db 163 GSGSVGGGGRGSGSVGSTGWTGTLFSENIIVTKNTROICDKGHLYKS-EVLNTG 221

QY 65 ASNDNHVFGYSTPWCYFPDNRFHCHFSRPRDQRLINNMGFRPKLNFNIQVKVTT 124

Db 222 DTAHRY-ATTPSYNFNQYSSHFSRPNQHLNDYERPRPAMTVRYNLIQKIMT 280

QY 125 NDGVTTIANN-LTSTVQVFSDBEYQLPYVLSAHQGLCPPADVFMIPOYGYLTN 179

Db 281 DGAMGTYYNDLTAGMHIFCDGHRVRYQHPMDQCMPELPSNIMELPOYAYIPAPISV 340

QY 180 ---NSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSVYASQSLDRL 232

Db 341 VDNNTVTVEEHLKGVPLVLENSDHEVLNNG-----RI 375

QY 233 MNPILDQYLYLNRTQ-----OSGSAQNKDLLFSRGP-----AGMSVQPKML 277

Db 376 YRIQYLMRLMDRKHIOHASDVOSTGQKQNLIIQKQPNKORFONAAALRTNM 435

QY 278 PGPCYRQORVSKTKDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSC 337

Db 436 SGP-----GIARGTNATLTQSGALVMTWNTGAD-----VSG 469

QY 338 VM-----IFGKE-----SAGASNTALDNVMTDEEIKATNPVAT 372

Db 470 VRAVVGSTDPVIGGGQPSDDLRLRYASASAAEQNPILN-----AAR 515

QY 373 ERFGTVANFQSSSTPATGCVHAMGALPGMWQDDEVYLOGPIWAKIPTDGHFHSPL 432

Db 516 HTFTRERATKLITSGNADGQYKEMWMLPQOMDSAPISRYNPVWVPRVNRKTLTLDQ 575

QY 433 MGFGFKHPPLQILIK--NTVPANPAPAEFSATKPFASITQYSGVSEIEMELOKENS 490

Db 576 DGSITPMGHPPTITIKLARIPVPGND-----SFLNITYYQGVSCVAVWEVERGT 626

QY 491 KRANPEVQYTSNVAKSAND-FTVDNNGLY 519

Db 627 KNMRPEYVHS--ATNMSVDAYTINNAGVY 653

RESULT 4

VCPV19

coat protein VP1 - parvovirus B19 (strain Au)

C/Species: parvovirus B19

A/Note: host Homo sapiens (man)

C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004

C/Accession: A24299

R/Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.

J. Virol. 58, 921-936, 1986

A/Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated from

A/Reference number: A24299; MUID:86200451; PMID:3701931

A/Accession: A24299

A/Molecule type: DNA

A/Residues: 1-781 <SHA>

A:Cross-references: UNIPROT:P07299; UNIPARC:UP10000127D7C; EMBL:M13178; NID:g333375; PID
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 16.8%; Score 487.5; DB 1; Length 781;

Best Local Similarity 27.4%; Pred. No. 8.3e-28;

Matches 143; Conservative 84; Mismatches 239; Indels 55; Gaps 16;

```
QY 10 ANNNEGADGVGNAS--GNMHCDSTWLGDRVITTTSTRTMALPTYNNHLYKQISSASTGASN 67
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 234 ABASTGAGGGGNSVSKMSSEGAIFSNVTCFSPQFLIPYDPEHHYKVFSPDASSCHN 293
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 68 D-----NHFGYSTPMGYFDPNRHCHSPDMDQRLINNMGFRPKALNFTLPMI 117
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 294 ASGKAECVCTISFIMGYSTPMRYLDENALNLFSPLEPHQLIENYGIAPDALTYTISI 353
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 118 QYKEVLT--TNDGVTIANNLSTVOVFSDESEYOLPVLGSAHQGCLPPPADVFMPOYG 175
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 354 AYKDVTDKTKGGV--QVTDSTIGRLCMLVDEHYKPYVLQGGQDTLAPELPIVYVPPQYA 412
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 176 YLTIAN--GSOAVG-----RSSFYCLEYFPSPQMLRTGNFTFSYTEEVPFSSVYAH 225
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 413 YLTGVVNTQGISGDSKSLASESAFYVLEHSSFOLLGTGTAAMS YKFPVPPEMLDEG 472
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 226 SGLDLKMLNPILIDLYLYINRTQNGSAONKDLFSRGS PAGMSVQPKMLPGCYRQO 285
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 473 SQHFYEMYNPL---YGSRLGVDPDTLGDPEKFRSL-----THEDHAIQPNFMFGPLVNSV 524
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 286 RVSKTTDNNNSFTWTGASKYNLNGRESLINRG--TAMASHKDEKFFPMGCVMLFGKE 344
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 525 STKSGSSNTGAKALTLGSLTGSTQNTRLSLRGPVVSQPYHMDTKYTGINALSHGQT 584
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 345 SAGASATYALDNV-----MITDEEIKATNPATERFGTVAVNFOSSSTDPATGDVHAM 397
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 585 TYG---NAEBKEVQGVGRPRNEKQLKQGLNHTY-----FNNKGTQVTDQIE-R 634
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 398 GALPGWVQDQVYVLOGPIWAKIPHTDGHFHS--PLMGFGFLKNPPQILLKNTPYVPA 456
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 635 PLWVGSVNMRRAHLYESQLSKIPNLDDSKFTQFALGGGGLHQPPOIFLK--ILPQSG 692
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 457 PAEFSATKPAFTTQVSTGVSEIEMEL-QKNSKRWMP 496
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 693 PIGGIKSMGITTLVQYAVGIMTVMTFKGPRKATGKRMNQ 733
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 5
VCPVVP
coat protein VP1 - porcine parvovirus
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
R:Kanz, A.I.; Mancius, J.U.; Diaz-Aroca, E.; Casael, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A:Title: Porcine parvovirus: DNA sequence and genome organization.
A:Reference number: A33302; MUID:90010964; PMID:2794971
A:Accession: B33302
A:Molecule type: DNA
A:Residues: 1-723 <RAN>
A:Cross-references: UNIPROT:P18546; UNIPARC:UP10000174964; EMBL:D00623
C:Genetics:
A:Initons: 10/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match 8.1%; Score 234.5; DB 1; Length 723;
Best Local Similarity 23.2%; Pred. No. 3.2e-09;
Matches 139; Conservative 83; Mismatches 241; Indels 137; Gaps 32;
```

```
QY 58 I-----SASTGAS--NDNHFGYSTPMGYFDPNRHCHSPDMDQRLINNMGFRPKALN 111
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 227 IHVLNSESAGSQVQADAHQWTPMSLIDANMGVFNPADQILNNMTEINLVFE 286
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 112 FKLFNIOYKEVLT--TNDGVTIANNLSTVOVFSDESEYOLPVLGSAHQGCLPPPADV 168
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 287 QALFNVVLTKTESATSPTKIYNNDLTASLMVLDNTNLTLPYTPAARSETLGFYPMIP 346
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 169 FMIPQGYL-----TLNNGSAVGRSS-----FYCLE-YFPQMLATGNKF 208
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 347 TKPTQRYRLSCIRMLNPPYTGQSQPNRLNTRNLHSDIMEFYTIENAVPIHLTLTGGEF 406
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 209 -TFSTEEVFPSSVYASQSILDR-----LNNPL--DOYLVLNRTQNGSAONK 257
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 407 STGIYHFDTKPL--KLTHSWQTRSLGLRPKLLTPTTEGDDPGSTPLAANTKRYHOTI 464
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 258 DLFSRGS PAGMSVQPKMLPGCYRQORVSKTKDNNNSFTWTGASKYNLNGRESLIN 317
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 465 NNSYTEAT---AIRP-----AQVGYNTPYMNEYNGGPF-----LT 498
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 318 P--GTAMASHKDEDEKFFPMGCVMLFGKESAGASTALD--NVMITDEEIK--ATNP- 369
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 499 PIVPTADQYUNDEEN-----GAIRFTMDYQHGLTSSQELERYTFYFNQ 543
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 370 ----VATERFQTV--VNFQSS-----TDPATG--DVHAMGALP-----GMWQDQRY 410
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 544 SKCGRAPKQGFNOQAPLNLNNTNNTLLPSDEIGKSNMNFNTNTTYGPLTALNNTAPV 603
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 411 YLOGPIWAKIPHTD--GHFPEPLMGFGFLK--NPPQILLKNTPYVPAVPAEFA--TKPA 466
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 604 FNGQIMDKELDTLKPRLH--VTAFFCYCKNPPQGLFVKIAP--NLTDFFNDSPPQ 657
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 467 SEITQYSTQGVSEIEMELQKNSKRWMPDQVYTSNYAKSANVDFTVNNGLYTEBRPIG 526
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 658 PRLITYSNMWMKGLTLFTAKKSSNMWNPICQHTT-----TAENIGHYI--PTNIG 706
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 6
A60006
coat protein VP1 - porcine parvovirus (strain 90HS)
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
R:Sakurai, M.; Nishimori, T.; Ushimi, C.; Nakajima, H.
Virus Res. 13, 79-86, 1989
A:Title: Nucleotide sequence of capsid protein gene of porcine parvovirus.
A:Reference number: A60006; MUID:89319168; PMID:2750278
A:Accession: A60006
A:Molecule type: DNA
A:Residues: 1-729 <SAK>
A:Cross-references: UNIPROT:P33484; UNIPARC:UP10000127C6E
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein; glycoprotein
F:151-729/Product: coat protein VP2 #status predicted <VP2>
F:172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #stact

Query Match 8.0%; Score 232.5; DB 1; Length 729;
Best Local Similarity 23.0%; Pred. No. 4.5e-09;
Matches 137; Conservative 86; Mismatches 245; Indels 127; Gaps 31;
```

```
QY 169 FMIPQGYL-----TLNNGSAVGRS-----SFYCLE-YFPGMLRTGNNF 208
DB 353 TKPQYRYVLLSCTRNLNPPYTGSOQITDSIQGLHSDIMFYTIENAVPIHLRTGDEF 412
QY 209 -TFSYTFEEVPHSSVYAHQSGLDR-----LNNPLI--DOYLYLNRTONSGSAQNK 257
DB 413 STGIYHFDTPRL--KLTHSQWTRSLGLPRKLLTEPTBEDQHPGTLPAANTRKGYHQT 470
QY 258 DLLFSRGPSPAGMSVOPKRWLPGPCYRQORVSKTKTDNNNSNFTMGASKYNLNGRESLIN 317
DB 471 NNSYTEAT-----AIRP-----AQGYNTPYNNFEYSNGGPF-----LT 504
QY 318 P--GTAMASHKDEDEKFFPMGSMIFGKESAGASNTALDNVMTDEBEIKATNP----- 369
DB 505 PIVPTADQYNDDE-----PNGAIRFTMGYQHGLTTS-----SGELERYTFNPGSKGR 554
QY 370 VATERFGTYA-VNFOSSS-----TDPATG--DVHAMGALP-----GVMQODRDVYLOGP 415
DB 555 APKQGFNOQAPLNLNTNNGTLPLPSDPIGKSNMHPMNTLNTYGPLTALNTAPVFPNGQ 614
QY 416 IMAKIPIPHD--GHFPPSLMGFGGLK--NPPQIILKNTPPVANPPEAFESA--TKFASFTQ 471
DB 615 IWDKELDTDLKPRHL--VTAPFVCKNPPQGLPFVKIAP--NLTDDEFNADSPQGPRIIT 668
QY 472 YSTGVSVSEIEMELQKENSKRNNPEVQYTSNYSKASANDFTVDNGLYTEPPPIG 526
DB 669 YSNFMWKGLTLFTAKRRSSNMNPIQOHTT-----TANIGNYI-PTNIG 712
```

RESULT 7

coat protein VP1 - porcine parvovirus (strain NADL-2)

N:Contains: coat protein VP2

C:Species: porcine parvovirus

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004

C:Accession: B33743; D48472

R:Vasudevacharya, U.; Basak, S.; Srinivas, R.V.; Compans, R.W.

Virology 173, 368-377, 1989

A:Title: Nucleotide sequence analysis of the capsid genes and the right-hand terminal pe

A:Reference number: A33743; MUID:90085785; PMID:2596019

A:Accession: B33743

A:Molecule type: DNA

A:Residues: 1-729 <VAS>

A:Cross-references: UNIPROT:P18546; UNIPARC:UPI0000127C70; GB:M32787; NID:G332983; PIDN:

R:Bergeron, U.; Menezes, U.; Tiljeen, P.

Virology 197, 86-98, 1993

A:Title: Genomic organization and mapping of transcription and translation products of p

A:Reference number: A48472; MUID:94025614; PMID:8212598

A:Accession: D48472

A:Molecule type: DNA

A:Residues: 11-729 <BER>

A:Cross-references: UNIPARC:UPI0000174965

A:Experimental source: NADL-2, ATCC VR-742

A>Note: sequence extracted from NCBI backbone (NCBIN:138789, NCBI:P138794)

C:Genetics:

A:introns: 10/1

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein; glycoprotein

F:151-729/Product: coat protein VP2 #status predicted <VP2>

F:32,172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 8.0%; Score 232.5; DB 1; Length 729;

Best Local Similarity 23.0%; Pred. No. 4,5e-09;

Matches 138; Conservative 85; Mismatches 240; Indels 137; Gaps 32;

```
QY 3 SGGGAPMADNNEGADGVGNAGNHCDS--TWIGD--RVITSTRMALPTYNHLYKO 57
DB 174 SGGGGG--GGGAGAGAGVSTGTGNNQTEFOYLGEBLVATIAHARLHLNMPREHETKR 232
QY 58 IS--SASTAS--NDNHVFGYSTPFWGYFDNRFCHFSPPDMQRLINNNMGFPKRLN 111
DB 233 IHLNSESGLVAGMVDADHTQMTPTWLSLDANAMGVWENPADWQLISNNMTEINLVSE 292
```

```
QY 112 FKLNIQKVEVT--TNDGVTIANNLTSVQVFSDEYOLPYVLGSAHQGCLPPFPADV 168
DB 293 QELFNVLKTIATESATSPPKIYNNDLTAISLVALDTNNTLPTPAAPRSETLGFPWLP 352
QY 169 FMIPQGYL-----TLNNGSAVGRS-----SFYCLE-YFPGMLRTGNNF 208
DB 353 TKPQYRYVLLSCTRNLNPPYTGSOQITDSIQGLHSDIMFYTIENAVPIHLRTGDEF 412
QY 209 -TFSYTFEEVPHSSVYAHQSGLDR-----LNNPLI--DOYLYLNRTONSGSAQNK 257
DB 413 STGIYHFDTPRL--KLTHSQWTRSLGLPRKLLTEPTBEDQHPGTLPAANTRKGYHQT 470
QY 258 DLLFSRGPSPAGMSVOPKRWLPGPCYRQORVSKTKTDNNNSNFTMGASKYNLNGRESLIN 317
DB 471 NNSYTEAT-----AIRP-----AQGYNTPYNNFEYSNGGPF-----LT 504
QY 318 P--GTAMASHKDEDEKFFPMGSMIFGKESAGASNTALD--NVMTDEBEIK--ATNP-- 369
DB 505 PIVPTADQYNDDEPN-----GAIPTMDYQHGLTTS--SGELERYTFNPG 549
QY 370 ----VATERFGTYA-VNFOSSS-----TDPATG--DVHAMGALP-----GVMQODRDV 410
DB 550 SKCGRAPKQGFNOQAPLNLNTNNGTLPLPSDPIGKSNMHPMNTLNTYGPLTALNTAPV 609
QY 411 YLOGPIWAKIPIHD--GHFPPSLMGFGGLK--NPPQIILKNTPPVANPPEAFESA--TKFA 466
DB 610 FPNQIWDKELDTDLKPRHL--VTAPFVCKNPPQGLPFVKIAP--NLTDDEFNADSPQ 663
QY 467 SFTQYSTGVSVSEIEMELQKENSKRNNPEVQYTSNYSKASANDFTVDNGLYTEPPPIG 526
DB 664 PRITTSNFMWKGLTLFTAKRRSSNMNPIQOHTT-----TANIGNYI-PTNIG 712
```

RESULT 8

coat protein VP1 - parvovirus Iu111

C:Species: parvovirus Iu111

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004

C:Accession: B44276

R:DiFloot, N.; Chen, K.C.; Bates, R.C.; Lederman, M.

Virology 192, 339-345, 1993

A:Title: The complete nucleotide sequence of parvovirus Iu111 and localization of a unig

A:Reference number: A44276; MUID:93297126; PMID:8517025

A:Accession: B44276

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-587 <DIF>

A:Cross-references: UNIPROT:P36310; UNIPARC:UPI0000127D7E; GB:M81888

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein; glycoprotein

F:49,90,220,304,371,503,511,514,539,571/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 7.9%; Score 229; DB 1; Length 587;

Best Local Similarity 22.7%; Pred. No. 5,9e-09;

Matches 133; Conservative 74; Mismatches 228; Indels 150; Gaps 28;

```
QY 3 SGGGAPMADNNEGADGVGNAGNHCDS--TWIGD--RVITSTRMAL--PTYNN--- 52
DB 30 SGGG-----GGGGGAGVSTGSDYNDQTHKKFLDGWVEITAVSTRVHLMPSSENYCR 83
QY 53 ---HLVQIISASTGANDNHVFGYSTPFWGYFDNRFCHFSPPDMQRLINNNMGFPKRL 109
DB 84 VRVINTNTDTGTAISMADDAHQW--TPWGLVDANAMGVWFPQSDWQYISNNMHIHLHS 142
QY 110 LNFLENIQKVEVT--TNDGVTIA--NNLSTVQVFSDEYOLPYVLGSAHQGCL--- 161
DB 143 LDGELFNVLKTIATESATSPPKIYNNDLTAAMVVALDSNNILPYTPALDNGETLGFYP 202
QY 162 --PEPPA-----DVFMIPQGYLTLNNGSAVGRSSSYCYCLEPPS--QMLATGN 206
DB 203 WKPTIPSPYRYRYSFCDRLNLSYTKDEAGTITDNGLASGLNSQFTLENTQRIINLRTGD 262
```


Db 638 YDPNATLISRIYTGTFPFMKGLTMRACKLRA-----NTTMNPVYQVSAB-----D 682
QY 515 NNGLY---TEPRPIGT-----RYLTRPL 534
Db 683 NGNSYMSVTKMLPTATGNMOSVPLITRPV 711

RESULT 11

VCPVIM

coat protein VP1 - minute virus of mice (strain WVM1)

C:Species: minute virus of mice, murine parvovirus

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004

C:Accession: B23008; B29510

R:Sahl, R.; McMaster, G.K.; Hirt, B.

Nucleic Acids Res. 13, 3617-3633, 1985

A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous

A:Reference number: A23008; MUID:85242059; PMID:3855242

A:Accession: B23008

A:Molecule type: DNA

A:Residues: 1-718 <SAH>

A:Cross-references: UNIPROT:P07302; UNIPARC:UPI0000174963; EMBL:X02481; NID:g60918; PIDN

R:Abstell, C.R.; Gardiner, E.M.; Tattersall, P.

J. Virol. 57, 656-669, 1986

A:Title: DNA sequence of the lymphotropic variant of minute virus of mice, WVM(1), and C

A:Reference number: A29510; MUID:86115415; PMID:3502703

A:Accession: B29510

A:Molecule type: DNA

A:Residues: 1-143, A, 145-718 <AST>

A:Cross-references: UNIPARC:UPI000127D6C; EMBL:M12032

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

Query Match 6.8%; Score 196.5; DB 1; Length 718;

Best Local Similarity 20.5%; Pred. No. 26-06;

Matches 122; Conservative 82; Mismatches 243; Indels 147; Gaps 25;

QY 3 SGGGAPMADNNEGADGVNAGSNMHCDS--TWLGD---RVITTSRTMALPTVNNHLLKQ 57
Db 161 SGGG-----GGGGGAGVSTGSDYNQTHYRFLSGGWETLALRLVHLNPKSENYCR 214
QY 58 I-----SSASTGASNDNHVFGYSTPMGWYDFNRFCHFSRPRDWQRIINNMGFRPRKL 110
Db 215 IYVHNTTDSYKGNMAKDHAHQIWTPMGLVDANMAGVWLQPSDWQYICNTMSQNLVSL 274
QY 111 NKLRIQYKQVETTTND---GVTTANNTSTVQYFSDSEYQLPVYLSAAGCLPPPPA 166
Db 275 DQEIFVVLKTYTEQDSGQAIKIYNNDLTACMVAADVSNILPYTPAANSMETLGFYFW 334
QY 167 DVFMIPQYGY-----LTINN-----GSAVGRSSFYCLEYFPS-OMLRTGN 206
Db 335 KPTIASPYRYFCVDRDLSTYENQGTIEHNVMTGPKMNSQFTIENTQOITLIRGCD 394
QY 207 NF-TFSYTEEEVFPSSVAHSQSLRLN--PLIDQYLYLNRTONGSSAQNKLFSR 263
Db 395 EFATGYVYDFDTPV--KLHTWQTRQLGQPLLSF---PEADTDACT-----LTAQ 442
QY 264 GSPAGNSVPRKWL-----PGCYRQQRVSKTIDNNNSNF 299
Db 443 GSRHGATQMEVNMVWSBAIRTPAQVFCOPHNDFEASRAGP-FAAPKVPADYTOGVDBEA 501
QY 300 TWTGASKYMLNGRESIINPGTAMASHKDEDF-----FPMGVMIFPKESAGASN 350
Db 502 NSGVSRYSGKQNGENMAAGPAPERITYTDETNFGSRDTRDGTIOSAPLVVPPPLANGILT 561
QY 351 TALDNWMTDEBEIKATNPATERFGTVANFQS--SSTDPAATGVHAGALPGWVODR 408
Db 562 NA-----NPIGTKN-----DIHFSNVNSYGPLASH-----PS 591
QY 409 DYLQGPPIAK--IHTDGHFHPSPLMGGFGLKAPPOILIKNTP--VPANPPAFSAT 463
Db 592 PVYPOQOIMDKELDHEKKRPLHITAPFV--CKNNAPGQMLVRLGPNLTQDYDPNGATLS 648

QY 464 KFASTFOYSTGVSEIEMELOKENSKRNPBEVQYTSNYAKSANDPTVNDNG 517
Db 649 RIYVYTGTFPFMKGLTMRACKLRA-----NTTMNPVYQ-----VSVEDNG 686

RESULT 12

VCPVME

coat protein VP1 - mink enteritis virus (strain Abashiri)

N:Contains: coat protein VP2

C:Species: mink enteritis virus, MEV

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C:Accession: B38350

R:Katsunari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shingai,

J. Gen. Virol. 72, 867-875, 1991

A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the

A:Reference number: A38350; MUID:9120123; PMID:2016597

A:Accession: B38350

A:Molecule type: DNA

A:Residues: 1-722 <KAR>

A:Cross-references: UNIPROT:P27437; UNIPARC:UPI0000174967; GB:D00765

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

F:113-722/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.7%; Score 194; DB 1; Length 722;

Best Local Similarity 20.7%; Pred. No. 3.1e-06;

Matches 136; Conservative 73; Mismatches 215; Indels 232; Gaps 32;

QY 5 GGAPMADNNEGADGVNAGSNMHCDSITLDRVYTTST-----RTVALPTVNN- 52
Db 148 GGGP-AVRNERATGSGNGSGGGGGGS--GGVGSTGTENNQTEFKLENGWEITNNS 204
QY 53 ---HL-----YKQI-----SSASTGASNDNHVFGYSTPMGWYDFNRFCHFSRPRDW 95
Db 205 RLVLHNPESBENYKRVVNNMDKTAVGNMALDDTHQIYTPMSLVDANMAGVFNFGDW 264
QY 96 QRLINNMGFRPKRLNFKLFIQYKVT---TNDGVTTANNTSTVQVSDSEYQLPVY 152
Db 265 QLIYNTMSLHVSFEQDIFVVLKTVSBSATQPTKYVNNDLTASIMVLDNSNTMPFT 324
QY 153 LGSAGGCL-----PPPPA-----DVFMIPQYGYLTNNGSAVAGS----- 189
Db 325 PAARSETLGFYPKKPIPTPMWRYFQWDRLLPSH-----TGTSGPTNIYHGTDPDD 378
QY 190 -SFYCLE-YPPSQMLRTGNPF--TF-----SYFEE-----VPFSSVAHSQSL 229
Db 379 VQFTTENSVPVHLRTGDEPATGTFFDCKPCRLTHTWQTRALGLPPLNLSPOSBA 438
QY 230 DR-----LMNPILIDYL---YLNRTONGS----- 251
Db 439 TNFGDIGVQDQKRGVYQWNTDYITATIRPAVEGYSAPYSPFBASTQGFPTIAG 498
QY 252 -GSAQNKLFSRSGSPAGMSVQPKMLPGCY--RQQRVSKTIDNNNSFTWTGASKY 307
Db 499 RGAQTDENQADDDPR-----YAGRGHQKQTTTGTGTPERFTY----- 538
QY 308 NLNGBESINGTAMASHKDEDFPFMSGVMINGXSAGASNTALD-NVMTDEBEIKA 366
Db 539 -----IAHQDT-----GRYPAGWIGNINPNLVTDNVLLP 570
QY 367 TNVPATERFGTVAVNFOSSSTDPAATGVHAGALPGWVODRDVYLOGPIYAKI PHTDGH 426
Db 571 TDPIG-----GKTGINY--TNFNYYGLTALNNP-----PYRPGQILMDKFPDDB- 615
QY 427 FHPSPLMGGFGLK-----NPPQILIKNTPVAN--PRAESATYFASFIT 470
Db 616 -----LKRRLHNAFVFCQNNCPQLFVKVAPNLNEBYDPRASANMSR-----IV 660
QY 471 QYSTGVSEIEMELOKENSKRNPBEVQYTSNYAKSANDPTVNDNGLYTEPRIG 526
Db 661 TYSDFWMKGLVFPKAKLRASHTNPIQMSIN-----VDNQFNLYL-PNNIG 705

A:Accession: B36608
A:Molecule type: DNA
A:Residues: 1-727 <MAR>
A:Cross-references: UNIPROT:P24840; UNIPARC:UPI0000127D56; GB:X55115
C:Superfamily: parovirus coat protein
C:Keywords: coat protein
F:144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.5%; Score 189; DB 1; Length 727;
Best Local Similarity 20.6%; Pred. No. 7,2e-06;
Matches 135; Conservative 73; Mismatches 216; Indels 232; Gaps 32;

```
QY 5 GGAPNADNNEGADVGNAGSNMHCSTWLGDRVITTTST-----RTMALPTVNN- 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 153 GGQP-AVNERATGSGNGSGGGGGG--CGVGISTGTFFNQTGKFLNGWEITANSS 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 53 ---HL-----YKQI-----SSASTGASNDNHFGISTPFGYFDNRFHCHFSPPDW 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 210 RLVLHMPESSENYKRVVNNMDKTAVKGMALDIIHVOIVTFPWSLVDANAMGVFNPQDW 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 96 QRLINNNQFRPKRLNFKLNFQVKEVT---TNDGVTIIANNLTSTVQVFSDEYQLPYV 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 270 QILIVNTMSBLHVSFEQELFNVLKTVBSATOPTKYVNDLTASLWALDSNNTMPFT 329
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 153 LGSANQGCT-----PPFPA-----DVFMIPOGYLTLNNGSQAVGRS----- 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 330 PAAMSEITGFPWKPTIPTPWRYFFQWDRILIPSH-----TGTSGTPTVYHGTDPDD 383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 -SFYCLE-YFPGQMLRTGNF--TF-----SYTPEE-----VPFHSSVAHSQSL 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 384 VQFYTIENSVPVHLRLRTGDEFATGFFFDCKPCRLTHWTQTRALGLPPFLNSLPOSEGA 443
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 DR-----LMNPLIDQYL--YLNRTQNS----- 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 444 TNFGDIGVQODKRGVTOGNTDYITBATIMRPAEVGISAPYSEFASTQGPFKTPIAAG 503
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 252 -GSAQNKDLLFSRGSPPAGMSVOPKMWLPQPCY--RQORVSKTKTDNNNSNFTWTGASKY 307
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 504 RGAQTDENQADGDR-----YAFGRQHQKTTTGTGTERPRTY----- 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 308 NINGRESIINPGTAMASHKDEDEKFFPMGSMVIFKESAGASNTALD-NVMITDEEETKA 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 544 -----IAHQDT-----GRYPEGDWIQININFLPVTNDNVLLP 575
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 367 TNPVATERGRTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGIWAKIPIHTDGH 426
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 576 TDPIG--GKTGINY--TNITFTYGPLTALNNVP-----PVYPNGQIMDKPDTD-- 620
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 427 FHPSPLMGSGFLK-----NPPQILIKNTVPVPAN--PPAEFSATKPAFRIT 470
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 621 -----LKPRLHVAPFVQCNCPGQLFVKVAPRLITNEYDPDASAMNSR-----IV 665
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 QYSTQVSVIEIMELQKENSKRMPDEVOYTSNYAKSANDVFTVDNNGLYTEPRPIG 526
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 666 TYSDFMWKGLVFKAKLRASHMTNPIDQMSIN-----VDNQFNYYV-PNNIG 710
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 23, 2005, 17:44:04
Job time : 20.8571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:25:34 ; Search time 102.857 Seconds
(without alignments)
3662.868 Million cell updates/sec

Title: US-10-696-282-17
Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLRLPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2906	100.0	736	2 Q9WB88_VVIRU	Q9WB88 adeno-ssoc
2	2892	99.5	736	2 Q6JC12_VVIRU	Q6JC12 adeno-ssoc
3	2885	99.3	736	2 Q6JC08_VVIRU	Q6JC08 adeno-ssoc
4	2884	99.2	736	2 Q6JC17_VVIRU	Q6JC17 adeno-ssoc
5	2883	99.2	736	2 Q6JC13_VVIRU	Q6JC13 adeno-ssoc
6	2863	98.5	736	2 Q6JC10_VVIRU	Q6JC10 adeno-ssoc
7	2536.5	87.3	735	2 Q6JC34_VVIRU	Q6JC34 adeno-ssoc
8	2535.5	87.3	735	2 Q6JC08_VVIRU	Q6JC08 adeno-ssoc
9	2532.5	87.1	735	2 Q6JC28_VVIRU	Q6JC28 adeno-ssoc
10	2530.5	87.1	735	2 Q6JB77_VVIRU	Q6JB77 adeno-ssoc
11	2530	87.1	736	2 Q6JC19_VVIRU	Q6JC19 adeno-ssoc
12	2529.5	87.0	735	2 Q6JB21_VVIRU	Q6JB21 adeno-ssoc
13	2528	87.0	736	2 Q6JC31_VVIRU	Q6JC31 adeno-ssoc
14	2524.5	86.9	735	2 Q6JC38_VVIRU	Q6JC38 adeno-ssoc
15	2524.5	86.9	735	2 Q6JC42_VVIRU	Q6JC42 adeno-ssoc
16	2520.5	86.7	735	2 Q6JC17_VVIRU	Q6JC17 adeno-ssoc
17	2517	86.6	734	2 Q6JC04_VVIRU	Q6JC04 adeno-ssoc
18	2515.5	86.6	735	2 Q6JC44_VVIRU	Q6JC44 adeno-ssoc
19	2514.5	86.5	735	2 Q6JC36_VVIRU	Q6JC36 adeno-ssoc
20	2513.5	86.5	735	2 Q6JB26_VVIRU	Q6JB26 adeno-ssoc
21	2511	86.4	734	2 Q6JC02_VVIRU	Q6JC02 adeno-ssoc
22	2506.5	86.3	735	2 Q6JC43_VVIRU	Q6JC43 adeno-ssoc
23	2502.5	86.1	735	2 Q6JB20_VVIRU	Q6JB20 adeno-ssoc
24	2500.5	86.0	735	2 Q6JB25_VVIRU	Q6JB25 adeno-ssoc
25	2499.5	86.0	735	2 Q6J0R4_VVIRU	Q6J0R4 adeno-ssoc
26	2493.5	85.8	735	2 Q6J0Q6_VVIRU	Q6J0Q6 adeno-ssoc
27	2493.5	85.8	735	2 Q6J0S0_VVIRU	Q6J0S0 adeno-ssoc
28	2493.5	85.8	735	2 Q6JC29_VVIRU	Q6JC29 adeno-ssoc
29	2492	85.8	734	2 Q6JC03_VVIRU	Q6JC03 adeno-ssoc
30	2491.5	85.7	735	2 Q6JC41_VVIRU	Q6JC41 adeno-ssoc
31	2490.5	85.7	735	2 Q6J0R6_VVIRU	Q6J0R6 adeno-ssoc

32	2490.5	85.7	735	2 Q6J0R8_VVIRU	Q6J0R8 adeno-ssoc
33	2488.5	85.6	735	2 Q6J0Q7_VVIRU	Q6J0Q7 adeno-ssoc
34	2488.5	85.6	735	2 Q6JB24_VVIRU	Q6JB24 adeno-ssoc
35	2488.5	85.6	735	2 Q6JC39_VVIRU	Q6JC39 adeno-ssoc
36	2484.5	85.5	735	2 Q6JC01_VVIRU	Q6JC01 adeno-ssoc
37	2483.5	85.5	735	2 Q6J0R0_VVIRU	Q6J0R0 adeno-ssoc
38	2483.5	85.5	735	2 Q6JC25_VVIRU	Q6JC25 adeno-ssoc
39	2483.5	85.5	735	2 Q6JC35_VVIRU	Q6JC35 adeno-ssoc
40	2481.5	85.4	533	2 Q92917_AAV2	Q92917 adeno-ssoc
41	2481.5	85.4	598	2 Q56653_AAV2	Q56653 adeno-ssoc
42	2481.5	85.4	735	2 Q56652_AAV2	Q56652 adeno-ssoc
43	2480.5	85.4	735	2 Q6JC24_VVIRU	Q6JC24 adeno-ssoc
44	2478.5	85.3	736	2 Q6JC24_VVIRU	Q6JC24 adeno-ssoc
45	2476.5	85.2	735	2 Q6JC33_VVIRU	Q6JC33 adeno-ssoc

ALIGNMENTS

RESULT 1					
Q9WB88_VVIRU					
ID Q9WB88_VVIRU PRELIMINARY;			PRT;	736 AA.	
AC Q9WB88;					
DT 01-NOV-1999 (TrEMBLrel. 12, Created)					
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)					
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)					
DE Capsid protein.					
OS Adeno-associated virus 1.					
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.					
OX NCBI_TaxID=85106;					
RN [1]					
RP NUCLEOTIDE SEQUENCE.					
RX MEDLINE=99214338; PubMed=10196295;					
RA Xiao W., Chirmule N., Berta S.C., McCullough B., Gao G., Wilson J.M.,					
RT "Gene therapy vectors based on adeno-associated virus type 1.";					
RU J. Virol. 73:3994-4003(1999).					
DR EMBL; AF063497; AAD27757.1; -; Genomic_DNA.					
DR SMR; Q9WB88; 217-736.					
DR GO; GO:0019028; C:Viral capsid; IEA.					
DR GO; GO:0005198; P:Structural molecule activity; IEA.					
DR InterPro; IPR001403; Parvo coat.					
DR Pfam; PF00740; Parvo coat.1.					
SQ SEQUENCE 736 AA; 81375 MW; CFABPB9H5CD0595 CRC64;					
Query Match	100.0%;	Score 2906;	DB 2;	Length 736;	
Best Local Similarity	100.0%;	Pred. No. 1.2e-203;			
Matches 534;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY 1 MASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTWALPTNNHLYKQISS 60					
DB 203 MASGGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTWALPTNNHLYKQISS 262					
QY 61 ASTGASNDNHYFGYSTPMGFEFNRFFCHSPRDMORLNNNGFRKRLNFKLFNIQVK 120					
DB 263 ASTGASNDNHYFGYSTPMGFEFNRFFCHSPRDMORLNNNGFRKRLNFKLFNIQVK 322					
QY 121 EYTTNDGVTIANNLTSTVQVFSDEYQLEPYVLSAHOGLPEPPADVFMIPOYGYTLN 180					
DB 323 EYTTNDGVTIANNLTSTVQVFSDEYQLEPYVLSAHOGLPEPPADVFMIPOYGYTLN 382					
QY 181 NSGQAVRSFFYCLEYRPSQMLRTGNFTSSYTFEEVPHSSVYHQSGLRIMNPLDOY 240					
DB 383 NSGQAVRSFFYCLEYRPSQMLRTGNFTSSYTFEEVPHSSVYHQSGLRIMNPLDOY 442					
QY 241 LYYLNRTONSGAONDLFFSRGSPAGMSVQPKMLPGPCYRQORVSKTKTDNNNSNFT 300					
DB 443 LYYLNRTONSGAONDLFFSRGSPAGMSVQPKMLPGPCYRQORVSKTKTDNNNSNFT 502					
QY 301 WTGASKYNLNGRESIINPGTAMASHKDDXKFFPMGCVMIFGKESAGASNTALDNWITD 360					
DB 503 WTGASKYNLNGRESIINPGTAMASHKDDXKFFPMGCVMIFGKESAGASNTALDNWITD 562					
QY 361 EETIKATNPVATERFGIVAVNFSSSTDPATGDVHAGALRGMWQDRDVLQGRPIAKI 420					

Db 563 EEEIKATNPVATERGTVAVNFOSSSTDPATCDVAMAGLPGMWQDDVYLQGIWAKI 622
Qy 421 PHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPAPAEFATKASFTIYOSTQOAVE 480
Db 623 PHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPAPAEFATKASFTIYOSTQOAVE 682
Qy 481 IEMELQKENSXKRWNPVEOYTSNYAKSANVDFTVDNNGLYTEBRPIGTRYLTRPL 534
Db 683 IEMELQKENSXKRWNPVEOYTSNYAKSANVDFTVDNNGLYTEBRPIGTRYLTRPL 736

RESULT 2

Q6JC12_VIRU PRELIMINARY; PRT; 736 AA.
AC Q6JC12;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TextID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues";
RL J. Virol. 78:6381-6388(2004).
DR EMBL: AY530607; SAS9292.1; -; Genomic_DNA.
DR SMR: Q6JC12; 217-736.
DR GO: GO:0019028; C:viral capsid; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81364 MW; F2D75A06662B6376 CRC64;

Query Match 99.5%; Score 2892; DB 2; Length 736;
Best Local Similarity 99.6%; Pred. No. 1.3e-202;
Matches 532; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MASGGGAPMADNNEGADGVNASGNMHCDSITWLGDRVITTSRTALPTNNHLYKOISS 60
Db 203 MASGGGAPMADNNEGADGVNASGNMHCDSITWLGDRVITTSRTALPTNNHLYKOISS 262
Qy 61 ASTGASNDNHFGYSTPFGYDFPNRPHCHFSPRDQRLINNMGFRPKLNFKLFNIQVK 120
Db 263 ASTGASNDNHFGYSTPFGYDFPNRPHCHFSPRDQRLINNMGFRPKLNFKLFNIQVK 322
Qy 121 EYTTNDGVTIANNLITSTVOVFSDBEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTIN 180
Db 323 EYTTNDGVTIANNLITSTVOVFSDBEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTIN 382
Qy 181 NGSQAVGRSSFYCLEYFPSONLRTGNFTFSYTFEEVPHSSYAHOSQIDRLMNPIDY 240
Db 383 NGSQAVGRSSFYCLEYFPSONLRTGNFTFSYTFEEVPHSSYAHOSQIDRLMNPIDY 442
Qy 241 LYTLNRTONOGSAGKNDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDNNNSNT 300
Db 443 LYTLNRTONOGSAGKNDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDNNNSNT 502
Qy 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 360
Db 503 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 562
Qy 361 EEEIKATNPVATERGTVAVNFOSSSTDPATGDVAMAGLPGMWQDDVYLQGIWAKI 420
Db 563 EEEIKATNPVATERGTVAVNFOSSSTDPATGDVAMAGLPGMWQDDVYLQGIWAKI 622
Qy 421 PHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPAPAEFATKASFTIYOSTQOAVE 480

Db 623 PHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPAPAEFATKASFTIYOSTQOAVE 682
Qy 481 IEMELQKENSXKRWNPVEOYTSNYAKSANVDFTVDNNGLYTEBRPIGTRYLTRPL 534
Db 683 IEMELQKENSXKRWNPVEOYTSNYAKSANVDFTVDNNGLYTEBRPIGTRYLTRPL 736

RESULT 3

Q6JC08_VIRU PRELIMINARY; PRT; 736 AA.
AC Q6JC08;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TextID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues";
RL J. Virol. 78:6381-6388(2004).
DR EMBL: AY530611; SAS9296.1; -; Genomic_DNA.
DR SMR: Q6JC08; 217-736.
DR GO: GO:0019028; C:viral capsid; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81317 MW; 373C7C40975CD9BD CRC64;

Query Match 99.3%; Score 2885; DB 2; Length 736;
Best Local Similarity 99.1%; Pred. No. 4.3e-202;
Matches 529; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MASGGGAPMADNNEGADGVNASGNMHCDSITWLGDRVITTSRTALPTNNHLYKOISS 60
Db 203 MASGGGAPMADNNEGADGVNASGNMHCDSITWLGDRVITTSRTALPTNNHLYKOISS 262
Qy 61 ASTGASNDNHFGYSTPFGYDFPNRPHCHFSPRDQRLINNMGFRPKLNFKLFNIQVK 120
Db 263 TSTGASNDNHFGYSTPFGYDFPNRPHCHFSPRDQRLINNMGFRPKLNFKLFNIQVK 322
Qy 121 EYTTNDGVTIANNLITSTVOVFSDBEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTIN 180
Db 323 EYTTNDGVTIANNLITSTVOVFSDBEYQLPYVLGSAHQGCLPPPADVFMIPOGYLTIN 382
Qy 181 NGSQAVGRSSFYCLEYFPSONLRTGNFTFSYTFEEVPHSSYAHOSQIDRLMNPIDY 240
Db 383 NGSQAVGRSSFYCLEYFPSONLRTGNFTFSYTFEEVPHSSYAHOSQIDRLMNPIDY 442
Qy 241 LYTLNRTONOGSAGKNDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDNNNSNT 300
Db 443 LYTLNRTONOGSAGKNDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDNNNSNT 502
Qy 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 360
Db 503 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 562
Qy 361 EEEIKATNPVATERGTVAVNFOSSSTDPATGDVAMAGLPGMWQDDVYLQGIWAKI 420
Db 563 EEEIKATNPVATERGTVAVNFOSSSTDPATGDVAMAGLPGMWQDDVYLQGIWAKI 622
Qy 421 PHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPAPAEFATKASFTIYOSTQOAVE 480
Db 623 PHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPAPAEFATKASFTIYOSTQOAVE 682
Qy 481 IEMELQKENSXKRWNPVEOYTSNYAKSANVDFTVDNNGLYTEBRPIGTRYLTRPL 534

Db 683 IEMWLOKENSKRNNPEVOYTSNVAKSANVDFVDDNGLYTEPRPIGRTYLRPL 736

RESULT 4

06J13_9VIRU PRELIMINARY; PRT; 736 AA.

AC 056137; 056137;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Capsid protein VPI.
OS Adeno-associated virus 6.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
NCBI_TaxID=68558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
DR EMBL; AF028704; AAB95450.1; -; Genomic_DNA.
DR SMR; 056137; 217-736.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo coat; 1.
SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;

Query Match 99.2%; Score 2884; DB 2; Length 736;

Best Local Similarity 99.1%; Pred. No. 5e-202; Matches 529; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MASGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTNNHLYQIIS 60
DB 203 MASGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTNNHLYQIIS 262
QY 61 ASTGASNDNHFGYSTPMGFDFPNRFCHFSPPDMQRLINNMGFPRKRLNFKLFNIQVK 120
DB 263 ASTGASNDNHFGYSTPMGFDFPNRFCHFSPPDMQRLINNMGFPRKRLNFKLFNIQVK 322
QY 121 EYTTNDGVTTIANLSTVQVPSDSEYQLPYVIGSAHQGLPPPADVPMIPOYGYLTIN 180
DB 323 EYTTNDGVTTIANLSTVQVPSDSEYQLPYVIGSAHQGLPPPADVPMIPOYGYLTIN 382
QY 181 NSQAVGRSSFYCLEFPPSQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLNPLIDQY 240
DB 383 NSQAVGRSSFYCLEFPPSQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLNPLIDQY 442
QY 241 LYYLNTONOSGSAQNKKDLFSRGSFAGMSVQPKMLPGPCYRQORVSKTKTDNNNSNFT 300
DB 443 LYYLNTONOSGSAQNKKDLFSRGSFAGMSVQPKMLPGPCYRQORVSKTKTDNNNSNFT 502
QY 301 WTGASKYNLNGRESIINPGTAMASHKDEBKFFPMGCVMI FGKESAGASTALDNNVITD 360
DB 503 WTGASKYNLNGRESIINPGTAMASHKDEBKFFPMGCVMI FGKESAGASTALDNNVITD 562
QY 361 EEEIKATNPVATERFGTVAANFOSSSTDPAITGVVHMGALPGMWQDRDYLLQGPITAKI 420
DB 563 EEEIKATNPVATERFGTVAANFOSSSTDPAITGVVHMGALPGMWQDRDYLLQGPITAKI 622
QY 421 PHTDGHFHPBLMGFGELKNPPOILLKNTPVANPAPAESATKFAFFITQYSTGQVSE 480
DB 623 PHTDGHFHPBLMGFGELKNPPOILLKNTPVANPAPAESATKFAFFITQYSTGQVSE 682
QY 481 IEMWLOKENSKRNNPEVOYTSNVAKSANVDFVDDNGLYTEPRPIGRTYLRPL 534
DB 683 IEMWLOKENSKRNNPEVOYTSNVAKSANVDFVDDNGLYTEPRPIGRTYLRPL 736

RESULT 5
06J13_9VIRU

ID 06J13_9VIRU PRELIMINARY; PRT; 737 AA.

AC 06J13; 06J13;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VPI.
CN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X., Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY510606; AAS99291.1; -; Genomic_DNA.
DR SMR; 06J13; 218-737.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo coat; 1.
SQ SEQUENCE 737 AA; 81332 MW; 94CED0F3AFD06E492 CRC64;

Query Match 99.2%; Score 2883; DB 2; Length 737;

Best Local Similarity 99.3%; Pred. No. 6e-202; Matches 530; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MASGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTNNHLYQIIS 60
DB 204 MASGGAPMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTNNHLYQIIS 263
QY 61 ASTGASNDNHFGYSTPMGFDFPNRFCHFSPPDMQRLINNMGFPRKRLNFKLFNIQVK 120
DB 264 ASTGASNDNHFGYSTPMGFDFPNRFCHFSPPDMQRLINNMGFPRKRLNFKLFNIQVK 323
QY 121 EYTTNDGVTTIANLSTVQVPSDSEYQLPYVIGSAHQGLPPPADVPMIPOYGYLTIN 180
DB 324 EYTTNDGVTTIANLSTVQVPSDSEYQLPYVIGSAHQGLPPPADVPMIPOYGYLTIN 383
QY 181 NSQAVGRSSFYCLEFPPSQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLNPLIDQY 240
DB 384 NSQAVGRSSFYCLEFPPSQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLNPLIDQY 443
QY 241 LYYLNTONOSGSAQNKKDLFSRGSFAGMSVQPKMLPGPCYRQORVSKTKTDNNNSNFT 300
DB 444 LYYLNTONOSGSAQNKKDLFSRGSFAGMSVQPKMLPGPCYRQORVSKTKTDNNNSNFT 503
QY 301 WTGASKYNLNGRESIINPGTAMASHKDEBKFFPMGCVMI FGKESAGASTALDNNVITD 360
DB 504 WTGASKYNLNGRESIINPGTAMASHKDEBKFFPMGCVMI FGKESAGASTALDNNVITD 563
QY 361 EEEIKATNPVATERFGTVAANFOSSSTDPAITGVVHMGALPGMWQDRDYLLQGPITAKI 420
DB 564 EEEIKATNPVATERFGTVAANFOSSSTDPAITGVVHMGALPGMWQDRDYLLQGPITAKI 623
QY 421 PHTDGHFHPBLMGFGELKNPPOILLKNTPVANPAPAESATKFAFFITQYSTGQVSE 480
DB 624 PHTDGHFHPBLMGFGELKNPPOILLKNTPVANPAPAESATKFAFFITQYSTGQVSE 683
QY 481 IEMWLOKENSKRNNPEVOYTSNVAKSANVDFVDDNGLYTEPRPIGRTYLRPL 534
DB 684 IEMWLOKENSKRNNPEVOYTSNVAKSANVDFVDDNGLYTEPRPIGRTYLRPL 737

RESULT 6
06J10_9VIRU
ID 06J10_9VIRU PRELIMINARY; PRT; 736 AA.
AC 06J10;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN NUCLEOTIDE SEQUENCE.
RP PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RX Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues";
RL J. Virol. 78:6381-6388(2004).
DR EMBL: AY530609; AAS9294.1; -; Genomic_DNA.
DR SMR: Q6UC10; 217-736.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81388 MW; BA9463E970028DF0 CRC64;

Query Match 98.5%; Score 2863; DB 2; Length 736;
Best Local Similarity 98.9%; Pred. No. 1.7e-200;
Matches 528; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MASGGAPMADNNGAGDGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 60
DB 203 MASGGAPMADNNGAGDGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPFGYDFPNRPHCHFSPRDMQRLINNNMGRPRKLNFKLFNIQVK 120
DB 263 ASTGASNDNHFGYSTPFGYDFPNRPHCHFSPRDMQRLINNNMGRPRKLNFKLFNIQVK 322
QY 121 ETTNDGVTTIANNLTSTVQVFSDBEYQLPYVLGSAHOGCLPPPADYFMIPOGYLTIN 180
DB 323 ETTNDGVTTIANNLTSTVQVFSDBEYQLPYVLGSAHOGCLPPPADYFMIPOGYLTIN 382
QY 181 NGSQAVGSSSFYCLEYFSPQMLRTGNNTFESTFEVPHSSYAHSSQSLDRMLNPLIDY 240
DB 382 NGSQAVGSSSFYCLEYFSPQMLRTGNNTFESTFEVPHSSYAHSSQSLDRMLNPLIDY 442
QY 241 LYYLNRTONGSQAONKDLFSRGSFAGSVQPKWMLPGPCYRQORVSKTKTDNNNSFT 300
DB 442 LYYLNRTONGSQAONKDLFSRGSFAGSVQPKWMLPGPCYRQORVSKTKTDNNNSFT 502
QY 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 360
DB 502 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 562
QY 361 EEBIKATNPVATERFGTAVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLOGPIMAKI 420
DB 562 EEBIKATNPVATERFGTAVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLOGPIMAKI 622
QY 421 PHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAFKFASFTIYOSTQGVSE 480
DB 622 PHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAFKFASFTIYOSTQGVSE 682
QY 481 IEMELQKENSKRMPNPEVQYTSNYAKSANDFTVDNNGLYTEBRPIGTRYLTRPL 534
DB 682 IEMELQKENSKRMPNPEVQYTSNYAKSANDFTVDNNGLYTEBRPIGTRYLTRPL 736

RESULT 7
Q6UC34_VIRU
ID Q6UC34_VIRU PRELIMINARY; PRT; 735 AA.
AC Q6UC34;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Capsid protein VP1.
OS Adeno-associated virus.
GN Name=cap;
OX NCBI_TaxID=272636;
RP NUCLEOTIDE SEQUENCE.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues";
RL J. Virol. 78:6381-6388(2004).
DR EMBL: AY530609; AAS9294.1; -; Genomic_DNA.
DR SMR: Q6UC10; 217-736.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81962 MW; 85DC69C5523D983 CRC64;

Query Match 87.3%; Score 2536.5; DB 2; Length 735;
Best Local Similarity 85.2%; Pred. No. 1.3e-176;
Matches 455; Conservative 34; Mismatches 44; Indels 1; Gaps 1;

QY 1 MASGGAPMADNNGAGDGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 60
DB 203 MASGGAPMADNNGAGDGVGNASGNMHCDSITWLGDRVITTSRTTALPTYNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPFGYDFPNRPHCHFSPRDMQRLINNNMGRPRKLNFKLFNIQVK 120
DB 263 OS-QASNDNHFGYSTPFGYDFPNRPHCHFSPRDMQRLINNNMGRPRKLNFKLFNIQVK 321
QY 121 ETTNDGVTTIANNLTSTVQVFSDBEYQLPYVLGSAHOGCLPPPADYFMIPOGYLTIN 180
DB 322 ETTNDGVTTIANNLTSTVQVFSDBEYQLPYVLGSAHOGCLPPPADYFMIPOGYLTIN 381
QY 181 NGSQAVGSSSFYCLEYFSPQMLRTGNNTFESTFEVPHSSYAHSSQSLDRMLNPLIDY 240
DB 382 NGSQAVGSSSFYCLEYFSPQMLRTGNNTFESTFEVPHSSYAHSSQSLDRMLNPLIDY 441
QY 241 LYYLNRTONGSQAONKDLFSRGSFAGSVQPKWMLPGPCYRQORVSKTKTDNNNSFT 300
DB 442 LYYLNRTONGSQAONKDLFSRGSFAGSVQPKWMLPGPCYRQORVSKTKTDNNNSFT 501
QY 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 360
DB 502 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTD 561
QY 361 EEBIKATNPVATERFGTAVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLOGPIMAKI 420
DB 562 EEBIKATNPVATERFGTAVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLOGPIMAKI 621
QY 421 PHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAFKFASFTIYOSTQGVSE 480
DB 622 PHTDGHFHSPLMGFGGLKNPPOILLIKNTVPANPAEFSAFKFASFTIYOSTQGVSE 681
QY 481 IEMELQKENSKRMPNPEVQYTSNYAKSANDFTVDNNGLYTEBRPIGTRYLTRPL 534
DB 682 IEMELQKENSKRMPNPEVQYTSNYAKSANDFTVDNNGLYTEBRPIGTRYLTRPL 735

RESULT 8
Q67008_VIRU
ID Q67008_VIRU PRELIMINARY; PRT; 735 AA.
AC Q67008;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RA Chen C.L., Jensen R.L., Schnepp B.C., Connell M.J., Bartlett J.S.,
RA Sfeera T.J., Shell R., Johnson P.R., Clark K.R.,
RT "Characterization of Adeno-Associated Viruses in Children."
RT Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY695376; AU05370.1; -; Genomic_DNA.
DR SMR: Q67008; 217-735.
DR GO: Q0019028; C:Vital capsid; IEA.
DR GO: Q0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo coat.
DR Pfam: PF00740; Parvo coat; 1.
DR SEQUENCE 735 AA; 81887 MW; 258BFB955F1C9925 CRC64;

Query Match 87.3%; Score 2535.5; DB 2; Length 735;
Best Local Similarity 85.2%; Pred. No. 1.5e-176;
Matches 455; Conservative 34; Mismatches 44; Indels 1; Gaps 1;

QY 1 MASGGAPADNNEGADGVNAGSNHCGSTWLGDRVITTSRTTMAIPYNNHLYQIIS 60
DB 203 MASGGAPADNNEGADGVNAGSNHCGSTWLGDRVITTSRTTMAIPYNNHLYQIIS 262
QY 61 ASTGASNDNHFGYSTPMGTFDNRPHCHFSPRDMQRLINNMGFPRKLNFLFNIOVK 120
DB 263 QS-GASNDNHFGYSTPMGTFDNRPHCHFSPRDMQRLINNMGFPRKLNFLFNIOVK 321
QY 121 EYTTNDGVTITANNLTSTVQVPSDSYQLPYVLGSAHQCLPPFPADVEMIPOYGYLTN 180
DB 322 EYTONDGTITANNLTSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVEMVPOYGYLTN 381
QY 181 NSQVAGRSFPCLETFPSCOMLRTGNNFTFSYTFEEVPHSSVAHSQSLDRLNPLIDQY 240
DB 382 NSQVAGRSFPCLETFPSCOMLRTGNNFTFSYTFEEVPHSSVAHSQSLDRLNPLIDQY 441
QY 241 LYLNTQNSGSAQNKDLFFSRGSPAGMSVOPKMLPGPCYRQORVSKTKTDNNNSNFT 300
DB 442 LYLNTQNSGSAQNKDLFFSRGSPAGMSVOPKMLPGPCYRQORVSKTKTDNNNSNFT 501
QY 301 WTGASKTYNLNGRESINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDNTWITD 360
DB 502 WTATATKXHLNGRSLVNPGRPMASHKODEDKFFPMGCVMI FGKESAGASNTALDNTWITD 561
QY 361 EBEIKATNPVATRFCTVAVNPOSSSTDPATGCVHAMGALPGVWMDRDVYLGPIWAKI 420
DB 562 EBEIRITNPVATRFCTVAVNPOSSSTDPATGCVHAMGALPGVWMDRDVYLGPIWAKI 621
QY 421 PHTDGHFHPSPILMGGFGLKMPPOQLIKNTVPANPAEFSATKFSFITQYSTGVSV 480
DB 622 PHTDGHFHPSPILMGGFGLKMPPOQLIKNTVPANPAEFSATKFSFITQYSTGVSV 681
QY 481 IEMELQENSKRNPNPEVQYTSNYSKASANDFTVDDNGLYTEPRIGTRYLTRPL 534
DB 682 IEMELQENSKRNPNPEVQYTSNYSKASANDFTVDDNGLYTEPRIGTRYLTRPL 735
RESULT 9
Q6JC28 GVIRU PRELIMINARY; PRT; 735 AA.
ID O6JC28 GVIRU PRELIMINARY;
AC O6JC28;
DT 05-JUL-2004 (Tremblere1.27, Created)
DT 05-JUL-2004 (Tremblere1.27, Last sequence update)
DE Capsid protein VPI.
GN Name=cap;
OC Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN NCLEBOTIDE SEQUENCE.
RP PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RX Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues." J. Virol. 78:6381-6388(2004).

DR EMBL: AY510591; AA599276.1; -; Genomic_DNA.
DR SMR: Q6JC28; 217-735.
DR GO: Q0019028; C:Vital capsid; IEA.
DR GO: Q0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo coat.
DR Pfam: PF00740; Parvo coat; 1.
DR SEQUENCE 735 AA; 81897 MW; 197066F8911FAE9B CRC64;

Query Match 87.1%; Score 2532.5; DB 2; Length 735;
Best Local Similarity 85.0%; Pred. No. 2.5e-176;
Matches 454; Conservative 34; Mismatches 45; Indels 1; Gaps 1;

QY 1 MASGGAPADNNEGADGVNAGSNHCGSTWLGDRVITTSRTTMAIPYNNHLYQIIS 60
DB 203 MASGGAPADNNEGADGVNAGSNHCGSTWLGDRVITTSRTTMAIPYNNHLYQIIS 262
QY 61 ASTGASNDNHFGYSTPMGTFDNRPHCHFSPRDMQRLINNMGFPRKLNFLFNIOVK 120
DB 263 QS-GASNDNHFGYSTPMGTFDNRPHCHFSPRDMQRLINNMGFPRKLNFLFNIOVK 321
QY 121 EYTTNDGVTITANNLTSTVQVPSDSYQLPYVLGSAHQCLPPFPADVEMIPOYGYLTN 180
DB 322 EYTONDGTITANNLTSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVEMVPOYGYLTN 381
QY 181 NSQVAGRSFPCLETFPSCOMLRTGNNFTFSYTFEEVPHSSVAHSQSLDRLNPLIDQY 240
DB 382 NSQVAGRSFPCLETFPSCOMLRTGNNFTFSYTFEEVPHSSVAHSQSLDRLNPLIDQY 441
QY 241 LYLNTQNSGSAQNKDLFFSRGSPAGMSVOPKMLPGPCYRQORVSKTKTDNNNSNFT 300
DB 442 LYLNTQNSGSAQNKDLFFSRGSPAGMSVOPKMLPGPCYRQORVSKTKTDNNNSNFT 501
QY 301 WTGASKTYNLNGRESINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDNTWITD 360
DB 502 WTATATKXHLNGRSLVNPGRPMASHKODEDKFFPMGCVMI FGKESAGASNTALDNTWITD 561
QY 361 EBEIKATNPVATRFCTVAVNPOSSSTDPATGCVHAMGALPGVWMDRDVYLGPIWAKI 420
DB 562 EBEIRITNPVATRFCTVAVNPOSSSTDPATGCVHAMGALPGVWMDRDVYLGPIWAKI 621
QY 421 PHTDGHFHPSPILMGGFGLKMPPOQLIKNTVPANPAEFSATKFSFITQYSTGVSV 480
DB 622 PHTDGHFHPSPILMGGFGLKMPPOQLIKNTVPANPAEFSATKFSFITQYSTGVSV 681
QY 481 IEMELQENSKRNPNPEVQYTSNYSKASANDFTVDDNGLYTEPRIGTRYLTRPL 534
DB 682 IEMELQENSKRNPNPEVQYTSNYSKASANDFTVDDNGLYTEPRIGTRYLTRPL 735
RESULT 10
Q6JB27 GVIRU PRELIMINARY; PRT; 735 AA.
ID O6JB27 GVIRU PRELIMINARY;
AC O6JB27;
DT 05-JUL-2004 (Tremblere1.27, Created)
DT 05-JUL-2004 (Tremblere1.27, Last sequence update)
DE Capsid protein VPI.
GN Name=cap;
OC Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN NCLEBOTIDE SEQUENCE.
RP PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RX Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues." J. Virol. 78:6381-6388(2004).
DR EMBL: AY530622; AA599307.1; -; Genomic_DNA.
DR SMR: Q6JB27; 217-735.
DR GO: Q0019028; C:Vital capsid; IEA.
DR GO: Q0005198; F:structural molecule activity; IEA.

DR InterPro: IPR001403; Parvo. Coat.
 DR Pfam: PF00740; Parvo coat; 1.
 DR SEQUENCE 735 AA; 81947 MW; EBA31AE904E0BF0A CRC64;

Query Match 87.1%; Score 2530.5; DB 2; Length 735;
 Best Local Similarity 84.8%; Pred. No. 3.6e-176;
 Matches 453; Conservative 35; Mismatches 45; Indels 1; Gaps 1;

```

QY 1 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTTSTRTMALPTYNHLYKQISS 60
DB 203 MASGGGAPMADNNEGADGVGNASGNMHCDSQMLGDRVITTTSTRTMALPTYNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPWCYGFDFNRFCHFSPRDQRLINNNGRRPKRLNFKLFNIOYK 120
DB 263 QS-GASNDNHFGYSTPWCYGFDFNRFCHFSPRDQRLINNNGRRPKRLNFKLFNIOYK 321
QY 121 EYTTNDGTTTANNLTSTYQVFSDEYQLPYVLGSAHQGCLPPPADVFPMIPOYGYLTIN 180
DB 322 EYTTNDGTTTANNLTSTYQVFTDSEYQLPYVLGSAHQGCLPPPADVFPMIPOYGYLTIN 381
QY 181 NSGQAVGRSSFYCLEYFSPQMLRTGNNTFSYTEEVEVPHSSYASHOSLDRMLNPLIDQY 240
DB 382 NSGQAVGRSSFYCLEYFSPQMLRTGNNTFSYTEEVEVPHSSYASHOSLDRMLNPLIDQY 441
QY 241 LYYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQORVSKTKTDNNNSNF 300
DB 442 LYYLNRTQNSGTLQOSRLTFSQAGPTNMSLQAKNMLPGPCYRQORLSQANDNNNSNF 501
QY 301 WTGASKTNLNGRESIINPGTAMASHKDEDEKFFPMGSMVIFKESAGASNTALDNVMTD 360
DB 502 WTAAKRYHLNGRDSLVNPGPAMASHKDEDEKFFPMHGNLIFGKOGTINADADLEVMITD 561
QY 361 DEEIKATNPVATERFGTVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLQGPIMAKI 420
DB 562 DEEIKATNPVATEIGTVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLQGPIMAKI 621
QY 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFAFTIOYSTGVSV 480
DB 622 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFAFTIOYSTGVSV 681
QY 481 IEMELQENSKRMNPEVQYTSNYSKASANDFTVDNNGLYTERPRIGTRYLTPL 534
DB 682 IEMELQENSKRMNPEIQYTSNYSKASANDFTVDNNGLYTERPRIGTRYLTPL 735

RESULT 11
ID 056139_VIRU PRELIMINARY; PRT; 736 AA.
AC 056139;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Capid protein VPI.
OS Adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxId=68742;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2."
RT J. Virol. 72:309-319(1998).
DR EMBL: AF028705; AAB95452.1; -; Genomic_DNA.
DR SMR: 056139; 217-736.
DR GO: GO:0019028; C:viral capid; IEA.
DR InterPro: IPR001403; Parvo coat.
DR Pfam: PF00740; Parvo coat; 1.
SQ SEQUENCE 736 AA; 81906 MW; DD52331AD5F0D70F CRC64;
```

Query Match 87.1%; Score 2530; DB 2; Length 736;
 Best Local Similarity 86.0%; Pred. No. 3.9e-176;

Matches 460; Conservative 27; Mismatches 46; Indels 2; Gaps 2;

```

QY 1 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITTTSTRTMALPTYNHLYKQISS 60
DB 203 MASGGGAPMADNNEGADGVGNASGNMHCDSQMLGDRVITTTSTRTMALPTYNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPWCYGFDFNRFCHFSPRDQRLINNNGRRPKRLNFKLFNIOYK 120
DB 263 QS-GASNDNHFGYSTPWCYGFDFNRFCHFSPRDQRLINNNGRRPKRLNFKLFNIOYK 321
QY 121 EYTTNDGTTTANNLTSTYQVFSDEYQLPYVLGSAHQGCLPPPADVFPMIPOYGYLTIN 180
DB 322 EYTTNDGTTTANNLTSTYQVFTDSEYQLPYVLGSAHQGCLPPPADVFPMIPOYGYLTIN 381
QY 181 NSGQAVGRSSFYCLEYFSPQMLRTGNNTFSYTEEVEVPHSSYASHOSLDRMLNPLIDQY 240
DB 382 NSGQAVGRSSFYCLEYFSPQMLRTGNNTFSYTEEVEVPHSSYASHOSLDRMLNPLIDQY 441
QY 241 LYYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQORVSKTKTDNNNSNF 299
DB 442 LYYLNRTQNSGTLQOSRLTFSQAGPTNMSLQAKNMLPGPCYRQORLSQANDNNNSNF 501
QY 300 WTGASKTNLNGRESIINPGTAMASHKDEDEKFFPMGSMVIFKESAGASNTALDNVMT 359
DB 502 WTAAKRYHLNGRDSLVNPGPAMASHKDEDEKFFPMHGNLIFGKGTINADADLEVMITD 561
QY 360 DEEIKATNPVATERFGTVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLQGPIMAK 419
DB 562 DEEIKATNPVATEIGTVAVNFOSSTDPATGDVHAMGALPGMWQODRDVYLQGPIMAK 621
QY 420 IHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFAFTIOYSTGVSV 479
DB 622 IHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFAFTIOYSTGVSV 681
QY 480 IEMELQENSKRMNPEVQYTSNYSKASANDFTVDNNGLYTERPRIGTRYLTPL 534
DB 682 IEMELQENSKRMNPEIQYTSNYSKASANDFTVDNNGLYTERPRIGTRYLTPL 736

RESULT 12
ID 06JBZ1_VIRU PRELIMINARY; PRT; 735 AA.
AC 06JBZ1;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Capid protein VPI.
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxId=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues."
RT J. Virol. 78:6381-6388(2004).
DR EMBL: AY530628; AAS99313.1; -; Genomic_DNA.
DR SMR: 06JBZ1; 217-735.
DR GO: GO:0019028; C:viral capid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo coat.
DR Pfam: PF00740; Parvo coat; 1.
SQ SEQUENCE 735 AA; 81868 MW; 1654ED6287F5474 CRC64;
```

Query Match 87.0%; Score 2529.5; DB 2; Length 735;
 Best Local Similarity 85.0%; Pred. No. 4.2e-176;
 Matches 454; Conservative 34; Mismatches 45; Indels 1; Gaps 1;

```
Db 203 MATGSGAPMADNNEGADGVNAGSNMHCDSQMLGDRVITTTSTRTMALPTYNHLYKQISS 262
Qy 61 ASTGASNDNHYFGYSTPMWGFDFNRPFHCHSPRDQRLINNNMGFRPKRLNFKLFINIOVK 120
Db 263 QS-GASNDNHYFGYSTPMWGFDFNRPFHCHSPRDQRLINNNMGFRPKRLNFKLFINIOVK 321
Qy 121 EVTTNDGVTITANNLSTVQVFSDESEYQLPYVLGSAHQGLCPPPADVFMIPQYGYLTJLN 180
Db 322 EVTQNDGVTITANNLSTVQVFTDSEYQLPYVLGSAHQGLCPPPADVFMVPOYGYLTJLN 381
Qy 181 NSQAVGRSSFFCYCLEYFSPQMLRTGNNTFFSYTFEEVPHSSVYAHOSQIDRLMNPIDQY 240
Db 382 NSQAVGRSSFFCYCLEYFSPQMLRTGNNTFFSYTFEEVPHSSVYAHOSQIDRLMNPIDQY 441
Qy 241 LYILNRTO-NOSGSAQNKDLFFSRGSPAGMSVQPKMLPGPCYRQQRVSFKTTDNNNSNF 300
Db 442 LYILNRTO-NOSGSAQNKDLFFSRGSPAGMSVQPKMLPGPCYRQQRVSFKTTDNNNSNF 501
Qy 301 WTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSMVIFGKESAGASNTALDNVMTD 360
Db 502 WTAAATKYHLNGRDSLVPNPGPAMASHKDEDEKFFPMHGTLLIFGKQTNADADLDNVMITD 561
Qy 361 EEEIKATNPVATERFETVAIVNFGSSSTDPAATGVNMGALPGMWQDRDVLVLOGPIMAKI 420
Db 562 EEEIRTTNPVATEQYGVYSNNLQNSNTGPTTGVNHOQALPGMWQDRDVLVLOGPIMAKI 621
Qy 421 PHTDGHFHPSPLMGFGGLKPNPQILIKNTVPANPPAEFSATKFASTTQYSTGQVSV 480
Db 622 PHTDGHFHPSPLMGFGGLKPNPQILIKNTVPANPPAEFSATKFASTTQYSTGQVSV 681
Qy 481 IEWELQKENSKRNPPEVOYTSNYAKSANDVTVDNNGLYTEPRPIGRTYLRPL 534
Db 682 IEWELQKENSKRNPPEVOYTSNYAKSANDVTVDNNGLYTEPRPIGRTYLRPL 735

RESULT 13
ID 065311.9VIRU PRELIMINARY; PRT; 736 AA.
AC 065311.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 3.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=46350;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=6266430; PubMed=8661429; DOI=10.1006/viro.1996.0367;
RA Muratatsu S., Mizukami H., Young N.S., Brown K.B.;
RT "Nucleotide sequencing and generation of an infectious clone of adeno-
associated virus 3.";
RL Virology 221:208-217(1996).
DR EMBL; U48704; AAC55049.1; -; Genomic_DNA.
SMR; 065311; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat; 1.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81660 MW; AFFLEF47B5C67A10 CRC64;

Query Match 87.0%; Score 2528; DB 2; Length 736;
Best Local Similarity 85.8%; Pred. No. 5.5e-176;
Matches 459; Conservative 28; Mismatches 46; Indels 2; Gaps 2;
```

```
Qy 121 EVTTNDGVTITANNLSTVQVFSDESEYQLPYVLGSAHQGLCPPPADVFMIPQYGYLTJLN 180
Db 322 EVTQNDGVTITANNLSTVQVFTDSEYQLPYVLGSAHQGLCPPPADVFMVPOYGYLTJLN 381
Qy 181 NSQAVGRSSFFCYCLEYFSPQMLRTGNNTFFSYTFEEVPHSSVYAHOSQIDRLMNPIDQY 240
Db 382 NSQAVGRSSFFCYCLEYFSPQMLRTGNNTFFSYTFEEVPHSSVYAHOSQIDRLMNPIDQY 441
Qy 241 LYILNRTO-NOSGSAQNKDLFFSRGSPAGMSVQPKMLPGPCYRQQRVSFKTTDNNNSNF 299
Db 442 LYILNRTO-NOSGSAQNKDLFFSRGSPAGMSVQPKMLPGPCYRQQRVSFKTTDNNNSNF 501
Qy 300 TWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSMVIFGKESAGASNTALDNVMT 359
Db 502 PWTAAASKYHLNGRDSLVPNPGPAMASHKDEDEKFFPMHGTLLIFGKQTNADADLDNVMIT 561
Qy 360 DEEIKATNPVATERFETVAIVNFGSSSTDPAATGVNMGALPGMWQDRDVLVLOGPIMAKI 419
Db 562 DEEIRTTNPVATEQYGVYANNLQSSNTAFTTGVNHOQALPGMWQDRDVLVLOGPIMAKI 621
Qy 420 IPHTDGHFHPSPLMGFGGLKPNPQILIKNTVPANPPAEFSATKFASTTQYSTGQVSV 479
Db 622 IPHTDGHFHPSPLMGFGGLKPNPQILIKNTVPANPPAEFSATKFASTTQYSTGQVSV 681
Qy 480 IEWELQKENSKRNPPEVOYTSNYAKSANDVTVDNNGLYTEPRPIGRTYLRPL 534
Db 682 IEWELQKENSKRNPPEVOYTSNYAKSANDVTVDNNGLYTEPRPIGRTYLRPL 736

RESULT 14
ID 06JC38.9VIRU PRELIMINARY; PRT; 735 AA.
AC 06JC38.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvirra M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530581; AAS99266.1; -; Genomic_DNA.
SMR; 06JC38; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat; 1.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81913 MW; 58131PDAEB024251 CRC64;

Query Match 86.9%; Score 2524.5; DB 2; Length 735;
Best Local Similarity 84.8%; Pred. No. 9.8e-176;
Matches 453; Conservative 34; Mismatches 46; Indels 1; Gaps 1;
```

```
Oy 181 NSGQAVGRSSFFCYCLEYFPSSOMLRTGNNTFFSYTFEEVFPFHSYAHQSGLDRLMPLIDQY 240
Db 382 NSGQAVGRSSFFCYCLEYFPSSOMLRTGNNTFFSYTFEEVFPFHSYAHQSGLDRLMPLIDQY 441
Oy 241 LYYLNRTONOGSQAONKDLFFSRGSPAGMSVQPKMWLPQPCYRQORVSKTKTDNNNSNFT 300
Db 442 LYYLNRTONOGSQAONKDLFFSRGSPAGMSVQPKMWLPQPCYRQORVSKTKTDNNNSNFT 501
Oy 301 WTGASKYLNLRGRESIIINPGTAMASHKODEDEKFFPMGSGVMTFGKESAGASNTALDNVMTD 360
Db 502 WTGASKYLNLRGRESIIINPGTAMASHKODEDEKFFPMGSGVMTFGKESAGASNTALDNVMTD 561
Oy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQODRDVYLQGPIMAKI 420
Db 562 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQODRDVYLQGPIMAKI 621
Oy 421 PHTDGHFHPSPLMGFGGLKPNPPQILIKNTVPANPAPAFSATAKFSFITYSTGQSVSE 480
Db 622 PHTDGHFHPSPLMGFGGLKPNPPQILIKNTVPANPAPAFSATAKFSFITYSTGQSVSE 681
Oy 481 IEMELQKENSKRMPBEIOYTSNYSKASANDFTVDNNGLYTEPRPIGTRYLTRNL 534
Db 682 IEMELQKENSKRMPBEIOYTSNYSKASANDFTVDNNGLYTEPRPIGTRYLTRNL 735

RESULT 15
O6JC42_9VIRU PRELIMINARY; PRT; 735 AA.
AC O6JC42;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
ON NCBI_TaxID=272636;
RX NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530577; AAS9262.1; -, Genomic_DNA.
DR SMR; Q6JC42; 217-735.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo coat; 1.
SQ SEQUENCE 735 AA; 82114 MM; 598748255744D5FB CRC64;

Query Match 86.9%; Score 2524.5; DB 2; Length 735;
Best Local Similarity 84.8%; Pred. No. 9, 8e-176;
Matches 453; Conservative 35; Mismatches 45; Indels 1; Gaps 1;

Oy 1 MASGQAPPADNNEGADGVGNAAGNHCSTWLGDRVITTSRTTMAALPTNNHLYKOISS 60
Db 203 MASGQAPPADNNEGADGVGNAAGNHCSTWLGDRVITTSRTTMAALPTNNHLYKOISS 262
Oy 61 ASTGASNDNHVFGYSTPMKGYPDNPRFHCHFSPPDQRLINNMMGFRPKLNFQIYQK 120
Db 263 AS-GASNDNHVFGYSTPMKGYPDNPRFHCHFSPPDQRLINNMMGFRPKLNFQIYQK 321
Oy 121 EYTTNDGVTITIANNLSTVQVFSDSSEYQHPYVLGSAHQCLPPPADVPMIPOYGYITLN 180
Db 322 EYTTNDGVTITIANNLSTVQVFSDSSEYQHPYVLGSAHQCLPPPADVPMIPOYGYITLN 381
Oy 181 NSGQAVGRSSFFCYCLEYFPSSOMLRTGNNTFFSYTFEEVFPFHSYAHQSGLDRLMPLIDQY 240
Db 382 NSGQAVGRSSFFCYCLEYFPSSOMLRTGNNTFFSYTFEEVFPFHSYAHQSGLDRLMPLIDQY 441
```

```
Oy 241 LYYLNRTONOGSQAONKDLFFSRGSPAGMSVQPKMWLPQPCYRQORVSKTKTDNNNSNFT 300
Db 442 LYYLNRTONOGSQAONKDLFFSRGSPAGMSVQPKMWLPQPCYRQORVSKTKTDNNNSNFT 501
Oy 301 WTGASKYLNLRGRESIIINPGTAMASHKODEDEKFFPMGSGVMTFGKESAGASNTALDNVMTD 360
Db 502 WTGASKYLNLRGRESIIINPGTAMASHKODEDEKFFPMGSGVMTFGKESAGASNTALDNVMTD 561
Oy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQODRDVYLQGPIMAKI 420
Db 562 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQODRDVYLQGPIMAKI 621
Oy 421 PHTDGHFHPSPLMGFGGLKPNPPQILIKNTVPANPAPAFSATAKFSFITYSTGQSVSE 480
Db 622 PHTDGHFHPSPLMGFGGLKPNPPQILIKNTVPANPAPAFSATAKFSFITYSTGQSVSE 681
Oy 481 IEMELQKENSKRMPBEIOYTSNYSKASANDFTVDNNGLYTEPRPIGTRYLTRNL 534
Db 682 IEMELQKENSKRMPBEIOYTSNYSKASANDFTVDNNGLYTEPRPIGTRYLTRNL 735
```

Search completed: November 23, 2005, 17:41:20
Job time : 104.857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 17:26:29 / Search time 24.2857 Seconds
(without alignments)
1817.892 Million cell updates/sec

Title: US-10-696-282-17

Perfect score: 2906
Sequence: 1 MASGGAPMADNNEGADGVG.....NNGLYTEPRPIGTIRYLRPL 534

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5.COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6.COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H.COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE.COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2906	100.0	534	2	US-09-807-802A-17 Sequence 17, Appl
2	2906	100.0	534	2	US-09-807-802A-15 Sequence 15, Appl
3	2906	100.0	736	2	US-09-807-802A-3 Sequence 3, Appl
4	2906	100.0	736	2	US-09-807-802A-13 Sequence 13, Appl
5	2481.5	85.4	533	2	US-10-038-972A-15 Sequence 15, Appl
6	2481.5	85.4	533	2	US-10-038-972A-14 Sequence 14, Appl
7	2481.5	85.4	735	2	US-09-321-589-1 Sequence 1, Appl
8	2481.5	85.4	735	2	US-10-293-478-1 Sequence 1, Appl
9	2481.5	85.4	735	2	US-10-038-972A-13 Sequence 13, Appl
10	1690.5	58.2	544	2	US-09-532-594B-18 Sequence 18, Appl
11	1690.5	58.2	544	2	US-09-532-594B-16 Sequence 16, Appl
12	1690.5	58.2	734	2	US-09-532-594B-4 Sequence 4, Appl
13	1665	57.3	532	2	US-09-533-427-6 Sequence 6, Appl
14	1665	57.3	588	2	US-09-533-427-5 Sequence 5, Appl
15	1665	57.3	724	2	US-09-533-427-4 Sequence 4, Appl
16	491.5	16.9	756	2	US-09-438-268-4 Sequence 4, Appl
17	486	16.7	554	2	US-10-187-253E-29 Sequence 29, Appl
18	486	16.7	554	2	US-10-187-253E-35 Sequence 35, Appl
19	486	16.7	781	2	US-10-187-253E-27 Sequence 27, Appl
20	479.5	16.5	781	2	US-08-856-841-22 Sequence 22, Appl
21	477	16.4	543	2	US-10-187-253E-33 Sequence 33, Appl
22	326	11.2	415	2	US-08-856-841-20 Sequence 20, Appl
23	318	10.9	395	2	US-08-856-841-13 Sequence 13, Appl
24	318	10.9	486	2	US-08-856-841-19 Sequence 19, Appl
25	317.5	10.9	500	2	US-08-856-841-16 Sequence 16, Appl
26	317.5	10.9	501	2	US-08-856-841-18 Sequence 18, Appl
27	304.5	10.5	264	2	US-08-856-841-14 Sequence 14, Appl

28	303	10.4	398	2	US-08-856-841-21	Sequence 21, Appl
29	209.5	7.2	579	6	5223424-13	Patent No. 5223424
30	183	6.3	584	2	US-09-022-949-2	Sequence 2, Appl
31	171	5.9	387	2	US-08-856-841-17	Sequence 17, Appl
32	132.5	4.6	489	2	US-10-376-397B-4	Sequence 4, Appl
33	119	4.1	1095	2	US-09-107-532A-3855	Sequence 3855, Ap
34	111.5	3.8	551	2	US-09-248-796A-23413	Sequence 23413, A
35	111	3.8	655	1	US-08-469-202-27	Sequence 27, Appl
36	111	3.8	655	1	US-08-484-434C-34	Sequence 34, Appl
37	111	3.8	655	2	US-09-384-361-34	Sequence 34, Appl
38	109	3.8	1186	1	US-08-485-568A-4	Sequence 4, Appl
39	109	3.8	1186	1	US-08-357-698-6	Sequence 6, Appl
40	109	3.8	1186	1	US-08-590-554A-4	Sequence 4, Appl
41	109	3.8	1186	1	US-09-184-223-4	Sequence 4, Appl
42	109	3.8	1186	4	PCT-US93-12682-6	Sequence 6, Appl
43	108.5	3.7	1637	2	US-10-172-502-14	Sequence 14, Appl
44	108	3.7	624	2	US-08-947-965-78	Sequence 78, Appl
45	108	3.7	655	1	US-08-469-202-28	Sequence 28, Appl

ALIGNMENTS

```

RESULT 1
US-09-807-802A-17
Sequence 17, Application US/09807802A
Patent No. 6759237
GENERAL INFORMATION:
APPLICANT: Xiao, Weidong
APPLICANT: Wilson, James M.
TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences.
FILE REFERENCE: GNVN.031USA
CURRENT APPLICATION NUMBER: US/09/807,802A
CURRENT FILING DATE: 2002-02-21
PRIORITY FILING DATE: 1998-11-05
PRIORITY APPLICATION NUMBER: PCT/US99/25694
PRIOR FILING DATE: 1999-11-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 17
LENGTH: 534
TYPE: PRT
ORGANISM: AAV-1
US-09-807-802A-17

Query Match      100.0%; Score 2906; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 9.6e-258;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MASGGAPMADNNEGADGVGNASGNMHCSTWLGDRVITTSRTWALPTYNNHLYKQISS 60
|||||
61 ASTGASNDNHFFCYSTPMGYPFDRFHCHFPSPDMQRLINNKGFPKALNFLEINIOVK 120
|||||
1 MASGGAPMADNNEGADGVGNASGNMHCSTWLGDRVITTSRTWALPTYNNHLYKQISS 60
|||||
61 ASTGASNDNHFFCYSTPMGYPFDRFHCHFPSPDMQRLINNKGFPKALNFLEINIOVK 120
|||||
121 EYTTNDGVTTIANNLSTYQVPSDSYOLPYVLSAHQGLPPPADVFVITPOYGYLTIN 180
|||||
121 EYTTNDGVTTIANNLSTYQVPSDSYOLPYVLSAHQGLPPPADVFVITPOYGYLTIN 180
|||||
181 NGSOAVGRSSPYCLEYFPQMLRTGNNFTFSYFEEVPPHSSVYAHSQSLDRLANPLIDY 240
|||||
181 NGSOAVGRSSPYCLEYFPQMLRTGNNFTFSYFEEVPPHSSVYAHSQSLDRLANPLIDY 240
|||||
181 NGSOAVGRSSPYCLEYFPQMLRTGNNFTFSYFEEVPPHSSVYAHSQSLDRLANPLIDY 240
|||||
241 LYLNTFONOSGAQNKDLLFSRGSFAGMSVOPKMLPGPCTRQGVSKTKDNNNSNFT 300
|||||
241 LYLNTFONOSGAQNKDLLFSRGSFAGMSVOPKMLPGPCTRQGVSKTKDNNNSNFT 300
|||||
301 WTGASVYNNNGRSIINPGTAMASHKDDDKFPFMSGWI FGESAGASNTALDNWITD 360
|||||

```

Db 301 WTGASKYVNLNGRESIIINPGTAMASHKDEDEKFFPMGSGVMI FKGESAGASNTALDVMITD 360
Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMGALPGMWQODDVLQGPIMAKI 420
Db 361 EEEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMGALPGMWQODDVLQGPIMAKI 420
Qy 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTITQYSTQGVSE 480
Db 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTITQYSTQGVSE 480
Qy 481 IEMELQKENS KRMNPEVOYTSNYAKSANDVFTV DNNGLYTEPRPIGTRYLTRPL 534
Db 481 IEMELQKENS KRMNPEVOYTSNYAKSANDVFTV DNNGLYTEPRPIGTRYLTRPL 534

RESULT 2

US-09-807-802A-15
/ Sequence 15, Application US/09807802A
/ Patent No. 6759237
/ GENERAL INFORMATION:
/ APPLICANT: Wilson, James M.
/ APPLICANT: Xiao, Weidong
/ TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
/ FILE REFERENCE: GNVN.031USA
/ CURRENT APPLICATION NUMBER: US/09/807,802A
/ PRIOR FILING DATE: 2002-02-21
/ PRIOR APPLICATION NUMBER: US 60/107,114
/ PRIOR FILING DATE: 1998-11-05
/ PRIOR APPLICATION NUMBER: PCT/US99/25694
/ PRIOR FILING DATE: 1999-11-02
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 15
/ LENGTH: 599
/ TYPE: PRT
/ ORGANISM: AAV-1
US-09-807-802A-15

Query Match 100.0%; Score 2906; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 1,1e-257;

Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITSTRITWALPTYNHLYKQISS 60
Db 66 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITSTRITWALPTYNHLYKQISS 125
Qy 61 ASTGASNDNHFGYSTPMGYDFNRFCHFSPRDQRLINNMGFRPKLNFKLFINIQVK 120
Db 126 ASTGASNDNHFGYSTPMGYDFNRFCHFSPRDQRLINNMGFRPKLNFKLFINIQVK 185
Qy 121 EYTTNDGVTITANNLTSTYQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQYGYLTIN 180
Db 186 EYTTNDGVTITANNLTSTYQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQYGYLTIN 245
Qy 181 NSQAVGRSSFYCLEYFSPQMLRTGNFTFSTFEVPHSSYAHQSGLDRLMNPILIDY 240
Db 246 NSQAVGRSSFYCLEYFSPQMLRTGNFTFSTFEVPHSSYAHQSGLDRLMNPILIDY 305
Qy 241 LYYLNRTONQSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTTDNNNSFT 300
Db 306 LYYLNRTONQSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTTDNNNSFT 365
Qy 301 WTGASKYVNLNGRESIIINPGTAMASHKDEDEKFFPMGSGVMI FKGESAGASNTALDVMITD 360
Db 366 WTGASKYVNLNGRESIIINPGTAMASHKDEDEKFFPMGSGVMI FKGESAGASNTALDVMITD 425
Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMGALPGMWQODDVLQGPIMAKI 420
Db 426 EEEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMGALPGMWQODDVLQGPIMAKI 485
Qy 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTITQYSTQGVSE 480
Db 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTITQYSTQGVSE 480

Db 486 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTITQYSTQGVSE 545
Qy 481 IEMELQKENS KRMNPEVOYTSNYAKSANDVFTV DNNGLYTEPRPIGTRYLTRPL 534
Db 546 IEMELQKENS KRMNPEVOYTSNYAKSANDVFTV DNNGLYTEPRPIGTRYLTRPL 599

RESULT 3

US-09-807-802A-3
/ Sequence 3, Application US/09807802A
/ Patent No. 6759237
/ GENERAL INFORMATION:
/ APPLICANT: Wilson, James M.
/ APPLICANT: Xiao, Weidong
/ TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
/ FILE REFERENCE: GNVN.031USA
/ CURRENT APPLICATION NUMBER: US/09/807,802A
/ PRIOR FILING DATE: 2002-02-21
/ PRIOR APPLICATION NUMBER: US 60/107,114
/ PRIOR FILING DATE: 1998-11-05
/ PRIOR APPLICATION NUMBER: PCT/US99/25694
/ PRIOR FILING DATE: 1999-11-02
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3
/ LENGTH: 736
/ TYPE: PRT
/ ORGANISM: AAV-1
US-09-807-802A-3

Query Match 100.0%; Score 2906; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 1,6e-257;

Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITSTRITWALPTYNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNASGNMHCDSITWLGDRVITSTRITWALPTYNHLYKQISS 262
Qy 61 ASTGASNDNHFGYSTPMGYDFNRFCHFSPRDQRLINNMGFRPKLNFKLFINIQVK 120
Db 263 ASTGASNDNHFGYSTPMGYDFNRFCHFSPRDQRLINNMGFRPKLNFKLFINIQVK 322
Qy 121 EYTTNDGVTITANNLTSTYQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQYGYLTIN 180
Db 323 EYTTNDGVTITANNLTSTYQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQYGYLTIN 382
Qy 181 NSQAVGRSSFYCLEYFSPQMLRTGNFTFSTFEVPHSSYAHQSGLDRLMNPILIDY 240
Db 383 NSQAVGRSSFYCLEYFSPQMLRTGNFTFSTFEVPHSSYAHQSGLDRLMNPILIDY 442
Qy 241 LYYLNRTONQSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTTDNNNSFT 300
Db 443 LYYLNRTONQSGAQNKDILFSRGS PAGMSVQPKMWLPGPCYRQORVSKTTDNNNSFT 502
Qy 301 WTGASKYVNLNGRESIIINPGTAMASHKDEDEKFFPMGSGVMI FKGESAGASNTALDVMITD 360
Db 503 WTGASKYVNLNGRESIIINPGTAMASHKDEDEKFFPMGSGVMI FKGESAGASNTALDVMITD 562
Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMGALPGMWQODDVLQGPIMAKI 420
Db 563 EEEIKATNPVATERFGTVAVNFQSSSTDPA TGDVHAMGALPGMWQODDVLQGPIMAKI 622
Qy 421 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTITQYSTQGVSE 480
Db 623 PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTITQYSTQGVSE 682
Qy 481 IEMELQKENS KRMNPEVOYTSNYAKSANDVFTV DNNGLYTEPRPIGTRYLTRPL 534
Db 683 IEMELQKENS KRMNPEVOYTSNYAKSANDVFTV DNNGLYTEPRPIGTRYLTRPL 736

RESULT 4

US-09-807-802A-13
; Sequence 13, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Xiao, Waidong
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPV.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-13

Query Match 100.0%; Score 2906; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 1,6e-257;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGAPMADNNEGADGVGNASGNHCDSTWIGDRVITTSRTWALPTYNHLYKQIS 60
DB 203 MASGGAPMADNNEGADGVGNASGNHCDSTWIGDRVITTSRTWALPTYNHLYKQIS 262

QY 61 ASTGASNDNHFGYSTPMWGFDPNRFCHFSPPDMQRLINNMGFRPKLNFLFNIOVK 120
DB 263 ASTGASNDNHFGYSTPMWGFDPNRFCHFSPPDMQRLINNMGFRPKLNFLFNIOVK 322

QY 121 EVTTNDGVTTIANNLSTVQVPSDSEYOLPYVLGSAHQCLPPFPADVPMIPOGYLTIN 180
DB 323 EVTTNDGVTTIANNLSTVQVPSDSEYOLPYVLGSAHQCLPPFPADVPMIPOGYLTIN 382

QY 181 NSGQAVGRSSFYCLEYFSPQMLRTGNNFTSYTFEEVPHSSVAHSOSLDRMLNPLIDQY 240
DB 383 NSGQAVGRSSFYCLEYFSPQMLRTGNNFTSYTFEEVPHSSVAHSOSLDRMLNPLIDQY 442

QY 241 LYYLNRTONQSGAQNKDLIFSRGSPAGMSVQPKMLPQPCYRQOVSKTKTDNNNSNFT 300
DB 443 LYYLNRTONQSGAQNKDLIFSRGSPAGMSVQPKMLPQPCYRQOVSKTKTDNNNSNFT 502

QY 301 WTGASKYNLNGRESIINPGTAMASHKODEKFFPMSCVMI FGKESAGASTALDNNWITD 360
DB 503 WTGASKYNLNGRESIINPGTAMASHKODEKFFPMSCVMI FGKESAGASTALDNNWITD 562

QY 361 EEEIKATNPVATERFGTVAANFQSSSTDPAATGVDHAMGALPGMWODRDVYLOGPIWAKI 420
DB 563 EEEIKATNPVATERFGTVAANFQSSSTDPAATGVDHAMGALPGMWODRDVYLOGPIWAKI 622

QY 421 PHTDGHFHPSPLMGGFGLKXPPQILIKNTVPANPAEFSATKFAFITQYSTGVSV 480
DB 623 PHTDGHFHPSPLMGGFGLKXPPQILIKNTVPANPAEFSATKFAFITQYSTGVSV 682

QY 481 IEWELQKENSKRANPEVOYTSNVAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 534
DB 683 IEWELQKENSKRANPEVOYTSNVAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 736

RESULT 5
US-10-038-972A-15
; Sequence 15, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038, 972A

CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 533
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP3 capsid protien
US-10-038-972A-15

Query Match 85.4%; Score 2481.5; DB 2; Length 533;
Best Local Similarity 83.3%; Pred. No. 8,3e-219;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGAPMADNNEGADGVGNASGNHCDSTWIGDRVITTSRTWALPTYNHLYKQIS 60
DB 1 MASGGAPMADNNEGADGVGNASGNHCDSTWIGDRVITTSRTWALPTYNHLYKQIS 60

QY 61 ASTGASNDNHFGYSTPMWGFDPNRFCHFSPPDMQRLINNMGFRPKLNFLFNIOVK 120
DB 61 ASTGASNDNHFGYSTPMWGFDPNRFCHFSPPDMQRLINNMGFRPKLNFLFNIOVK 119

QY 121 EVTTNDGVTTIANNLSTVQVPSDSEYOLPYVLGSAHQCLPPFPADVPMIPOGYLTIN 180
DB 120 EVTTNDGVTTIANNLSTVQVPSDSEYOLPYVLGSAHQCLPPFPADVPMIPOGYLTIN 179

QY 181 NSGQAVGRSSFYCLEYFSPQMLRTGNNFTSYTFEEVPHSSVAHSOSLDRMLNPLIDQY 240
DB 180 NSGQAVGRSSFYCLEYFSPQMLRTGNNFTSYTFEEVPHSSVAHSOSLDRMLNPLIDQY 239

QY 241 LYYLNRTONQSGAQNKDLIFSRGSPAGMSVQPKMLPQPCYRQOVSKTKTDNNNSNFT 300
DB 240 LYYLNRTONQSGAQNKDLIFSRGSPAGMSVQPKMLPQPCYRQOVSKTKTDNNNSNFT 299

QY 301 WTGASKYNLNGRESIINPGTAMASHKODEKFFPMSCVMI FGKESAGASTALDNNWITD 360
DB 300 WTGASKYNLNGRESIINPGTAMASHKODEKFFPMSCVMI FGKESAGASTALDNNWITD 359

QY 361 EEEIKATNPVATERFGTVAANFQSSSTDPAATGVDHAMGALPGMWODRDVYLOGPIWAKI 420
DB 360 EEEIKATNPVATERFGTVAANFQSSSTDPAATGVDHAMGALPGMWODRDVYLOGPIWAKI 419

QY 421 PHTDGHFHPSPLMGGFGLKXPPQILIKNTVPANPAEFSATKFAFITQYSTGVSV 480
DB 420 PHTDGHFHPSPLMGGFGLKXPPQILIKNTVPANPAEFSATKFAFITQYSTGVSV 479

QY 481 IEWELQKENSKRANPEVOYTSNVAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 534
DB 480 IEWELQKENSKRANPEVOYTSNVAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 533

RESULT 6
US-10-038-972A-14
; Sequence 14, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038, 972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 598
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP2 capsid protien
US-10-038-972A-14

Query Match 85.4%; Score 2481.5; DB 2; Length 598;

Best Local Similarity 83.3%; Pred. No. 9.9e-219;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGAPMADNNEGADGVGNASGNHCDSTWLGDRVITTSRTTALPTYNHLYKQISS 60
DB 66 MATGSGAPMADNNEGADGVGNASGNHCDSTWMDGRVITTSRTTALPTYNHLYKQISS 125
QY 61 ASTGASNDNHFGYSTPWCYFDPNRFCHFSFPRDQRLINNMGFRPKLNFKLNIQVK 120
DB 126 QGASNDNHFGYSTPWCYFDPNRFCHFSFPRDQRLINNMGFRPKLNFKLNIQVK 184
QY 121 EYTNNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQGYLTLN 180
DB 185 EYTNNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQGYLTLN 244
QY 181 NGSQAVGRSSFYCLEYFSPQMLRTGNNTFSTFEEDVPHSSYAHQSGLDRMLNPLIDQY 240
DB 245 NGSQAVGRSSFYCLEYFSPQMLRTGNNTFSTFEEDVPHSSYAHQSGLDRMLNPLIDQY 304
QY 241 LYTLNRTONOGSGAQNKDILFSRGSFAGMSVOPKMWLPQPCYRQORVSKTXDNNNSFT 300
DB 305 LYTLNRTONOGSGAQNKDILFSRGSFAGMSVOPKMWLPQPCYRQORVSKTXDNNNSFT 364
QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMTFGKESAGASNTALDNVMTD 360
DB 365 WTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMTFGKESAGASNTALDNVMTD 424
QY 361 EEEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMALPGMWODRDVYLQGPIMAKI 420
DB 425 EEEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMALPGMWODRDVYLQGPIMAKI 484
QY 421 PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSAATKFSFTQYSTGVSE 480
DB 485 PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSAATKFSFTQYSTGVSE 544
QY 481 IEMELQKENSXKRNWPEVQYTSNYAKSANDFTVNNGLYTEPRIGTRYLTRPL 534
DB 545 IEMELQKENSXKRNWPEVQYTSNYAKSANDFTVNNGLYTEPRIGTRYLTRPL 598

RESULT 7
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SATIL D.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match 85.4%; Score 2481.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 1.4e-218;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

DB 322 EYTNNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQGYLTLN 381
QY 181 NGSQAVGRSSFYCLEYFSPQMLRTGNNTFSTFEEDVPHSSYAHQSGLDRMLNPLIDQY 240
DB 382 NGSQAVGRSSFYCLEYFSPQMLRTGNNTFSTFEEDVPHSSYAHQSGLDRMLNPLIDQY 441
QY 241 LYTLNRTONOGSGAQNKDILFSRGSFAGMSVOPKMWLPQPCYRQORVSKTXDNNNSFT 300
DB 442 LYTLNRTONOGSGAQNKDILFSRGSFAGMSVOPKMWLPQPCYRQORVSKTXDNNNSFT 501
QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMTFGKESAGASNTALDNVMTD 360
DB 502 WTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMTFGKESAGASNTALDNVMTD 561
QY 361 EEEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMALPGMWODRDVYLQGPIMAKI 420
DB 562 EEEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMALPGMWODRDVYLQGPIMAKI 621
QY 421 PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSAATKFSFTQYSTGVSE 480
DB 622 PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSAATKFSFTQYSTGVSE 681
QY 481 IEMELQKENSXKRNWPEVQYTSNYAKSANDFTVNNGLYTEPRIGTRYLTRPL 534
DB 682 IEMELQKENSXKRNWPEVQYTSNYAKSANDFTVNNGLYTEPRIGTRYLTRPL 735

RESULT 8
US-10-293-478-1
; Sequence 1, Application US/10293478
; Patent No. 6733757
; GENERAL INFORMATION:
; APPLICANT: PATEL, SATIL D.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 85.4%; Score 2481.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 1.4e-218;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

RESULT 11

```
US-09-532-594B-16
; Sequence 16, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kofin, Robert M.
; APPLICANT: Safier, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532.594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc_feature
; OTHER INFORMATION: AAV4 capsid protein VP2
US-09-532-594B-16
```

```
Query Match      58.2%; Score 1690.5; DB 2; Length 598;
Best Local Similarity 59.4%; Pred. No. 3.6e-146;
Matches 325; Conservative 65; Mismatches 132; Indels 25; Gaps 9;
```

```
QY 2 ASGGAPMADNNEGADGVNAGSNHCDSTWLGDRVITTSRTWALPTYNHNLKYQISGA 61
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 63 AAGAGAAV-EGGQADGADGVNAGSDHCDSTWSEGHVTTTSTRTWALPTYNHNLKYRLGE- 120
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 STGASNDNHFGYSTPMGYFDNFRFHCHFSPRDMORLIINNMGFRPKLNFLENIQYKE 121
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 121 ---SLQSTNYNGFSTPMGYFDNFRFHCHFSPRDMORLIINNMGFRPKARVXIFNIQYKE 177
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 VTTNDGVTIANNTSTVQVPSDSEYOLPYVLGSAHQGCLPPFADVFMIPOYGY---LT 178
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 178 VTTNGETTVANNLTSTVQIFADSSYELPYVMDAGQSGSLPPFPNDVFMVQYCGLYT 237
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 179 LNNGSAVGRSSFYCLEYFPQSOMLRTGNFTFSYFEEVPRHSSVYAHQSGLDRLMNP 238
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 238 GNTSQQOTDRNAFYCLEYFPQSOMLRTGNFTFSYFEEVPRHSSVYAHQSGLDRLMNP 297
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 239 QYLYLNFRTON---QSGSAQNKDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDN 294
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 298 QYLMGLQSTTTGTLINAGTATTN---FTKLRPTNFSNPKKMWLPGPSIKQOGFSKTA--N 352
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 295 NNSNFTWTGAS---KY---NLNGREBSIINGTAMASHKDEDEKFFPMGVMIFGKESAG 347
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 353 QNYKIPATGSDSLIKYETHSTLDGRWSALTGPMPATAGPADSK--FNSSQLIFAGPKONG 411
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 348 ASNTALDNWMTDEBEIKATNPVATERGTVAVNFQSSSTDPAIDGVAMGALPGMWOD 407
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 412 NTATVPGLITFTSEBELATNATDTDMWGNNLPGDQSSNSNLPYDRLTALGAVPGMWQN 471
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 408 RDVYLQGPWAKIPHTDGHFSPPLMGFGGLKAPPPQILIKNTVPANPAPAEFSATKAS 467
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 472 RDIYVQGIWAKIPIHTDGHFSPPLIGFGGLKHPPPQIFIKNTVPANPAPATTFSSTPVNS 531
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 468 FITYSTGVQVSEIEMELQKNSKRNPEVOYTSNVAASAVDFTVDNNGLYTEPRPIGT 527
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 532 FITYSTGVQVQIDMEIQKRSKRNPEVOFTSNYQGNLSLMAPDAAAGKYTEBRAIGT 591
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 528 RYLTRPL 534
   |||
DB 592 RYLTHHL 598
```

RESULT 12

```
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kofin, Robert M.
; APPLICANT: Safier, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532.594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc_feature
; OTHER INFORMATION: VPI
US-09-532-594B-4
```

```
Query Match      58.2%; Score 1690.5; DB 2; Length 734;
Best Local Similarity 59.4%; Pred. No. 4.9e-146;
Matches 325; Conservative 65; Mismatches 132; Indels 25; Gaps 9;
```

```
QY 2 ASGGAPMADNNEGADGVNAGSNHCDSTWLGDRVITTSRTWALPTYNHNLKYQISGA 61
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 199 AAGAGAAV-EGGQADGADGVNAGSDHCDSTWSEGHVTTTSTRTWALPTYNHNLKYRLGE- 256
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 STGASNDNHFGYSTPMGYFDNFRFHCHFSPRDMORLIINNMGFRPKLNFLENIQYKE 121
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 257 ---SLQSTNYNGFSTPMGYFDNFRFHCHFSPRDMORLIINNMGFRPKARVXIFNIQYKE 313
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 VTTNDGVTIANNTSTVQVPSDSEYOLPYVLGSAHQGCLPPFADVFMIPOYGY---LT 178
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 314 VTTNGETTVANNLTSTVQIFADSSYELPYVMDAGQSGSLPPFPNDVFMVQYCGLYT 373
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 179 LNNGSAVGRSSFYCLEYFPQSOMLRTGNFTFSYFEEVPRHSSVYAHQSGLDRLMNP 238
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 374 GNTSQQOTDRNAFYCLEYFPQSOMLRTGNFTFSYFEEVPRHSSVYAHQSGLDRLMNP 433
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 239 QYLYLNFRTON---QSGSAQNKDLLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTKTDN 294
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 434 QYLMGLQSTTTGTLINAGTATTN---FTKLRPTNFSNPKKMWLPGPSIKQOGFSKTA--N 488
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 295 NNSNFTWTGAS---KY---NLNGREBSIINGTAMASHKDEDEKFFPMGVMIFGKESAG 347
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 489 QNYKIPATGSDSLIKYETHSTLDGRWSALTGPMPATAGPADSK--FNSSQLIFAGPKONG 547
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 348 ASNTALDNWMTDEBEIKATNPVATERGTVAVNFQSSSTDPAIDGVAMGALPGMWOD 407
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 548 NTATVPGLITFTSEBELATNATDTDMWGNNLPGDQSSNSNLPYDRLTALGAVPGMWQN 607
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 408 RDVYLQGPWAKIPHTDGHFSPPLMGFGGLKAPPPQILIKNTVPANPAPAEFSATKAS 467
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 608 RDIYVQGIWAKIPIHTDGHFSPPLIGFGGLKHPPPQIFIKNTVPANPAPATTFSSTPVNS 667
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 468 FITYSTGVQVSEIEMELQKNSKRNPEVOYTSNVAASAVDFTVDNNGLYTEPRPIGT 527
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 668 FITYSTGVQVQIDMEIQKRSKRNPEVOFTSNYQGNLSLMAPDAAAGKYTEBRAIGT 727
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 528 RYLTRPL 534
   |||
DB 728 RYLTHHL 734
```

RESULT 13
US-09-533-427-6

```

: Sequence 6 Application US/09533427
: Patent No. 6855314
: GENERAL INFORMATION:
: APPLICANT: Chiorini, John
: APPLICANT: Kotlin, Robert M.
: APPLICANT: Safer, Brian
: APPLICANT: Davidson, Elizabeth
: APPLICANT: Zahner, Joseph
: TITLE OR INVENTION: JAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
: FILE REFERENCE: 14014.0323U2
: CURRENT APPLICATION NUMBER: US/09/533,427
: CURRENT FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 532
: TYPE: PR1
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
: US-09-533-427-6

```

```

Query Match 57.3%; Score 1665; DB 2; Length 532;
Best Local Similarity 58.9%; Pred. No. 6, 5e-144;
Matches 318; Conservative 54; Mismatches 154; Indels 14; Gaps 5

QY      1 MASGGGADPMADNNEGADGVNAGSNWCHDSITWLGDVRVITTSRTTALPTTYNNHLYKQISS 60
DB      1 MSAGGGGGLGDNNQCADGVNAGSNWCHDSITWNGRIVVTKSTRTWLPSTYNNHYREIKS 60
QY      61 ASTGASNNHYGYSCTPMWGYPDPENRPHCHGFSRDMORLNNMWGPPKRLNPLFLPIQVK 120
DB      61 GSVDSNNANAFGYSTPMWGYPDPENFPHSHMSRDMORLNNMWGFRKRLRYKIFNIQVK 120
QY      121 EYTTNDGYTTIANNLITSTQVPSDSEYQLPYVLGASAHQCLRPFPADVEMIPQYGLTLN 180
DB      121 EYTVQDSITTIANNLITSTVQVFTDDYQLPYVVGNTGECCLRPFPQVFTLPOGYATLN 180
QY      181 --NSGQAVGRSSFFYCLEYFSPQMLRTGNNFTSSYTFEEVPPFHSSYVHSGSLDRMLPLD 238
DB      181 RDNTNPTERSSEFFCLEYFSPSKRLRTGNNFEFTYNEEVPFHSSFPASQNLFTLANPLVD 240
QY      239 QYLYLNRQNGSGSAQNDDLFSRGS PAGMSVQP -KMLPGPCYQQRVSKTKTNNNS 297
DB      241 QYLYFVSTNNNGYQVFNKL-----AGRYANTYKMFPPGPMGTQGMNLGSGNRAS 293
QY      298 NETTWGASKYNNNGRESIINPGTAMASHKDEDEKFFPMSGVMI FGKESAGASNTAL--D 354
DB      294 VSAFPTNRNMELEGASVQVPQPNGMTNNLQGSNTVALENTMIFNSGPANPGTTATYLEG 353
QY      355 NVMTIDEDEETKATNPYATREFGTVAVNPFSSSDPDTGTVHMGALPGMWQDRDVLQ 414
DB      354 NMLITSESETOGVNRYAVVVGQMAATNNOSSTTAAPTGTYNLOEIVPGSVMERDVLQ 413
QY      415 PIWAKIPIHTDGFHPSPLEMGFGELKQKPPQILIKNTPVANPAPESATKFAFITQYST 474
DB      414 PIWAKIPIETGALFHBPAPMGFGELKRPPIEMMLIKNTPVGN -ITSSDVPVSSFIQYST 472
QY      475 GGVSYVEIEMELQKENSKRNNPEVOYTSNAYKASANDVFTVDNNGLYTEPRPIGRTYLRPL 534
DB      473 GGVTEMEMELKKENSKRNNPEIQYNTNVDQFVDVAPADSTGEYRTTRPIGRTYLRPL 532

RESULT 14
US-09-533-427-5
; Sequence 5, Application US/09533427
; Patent No. 6855314
; GENERAL INFORMATION:
; APPLICANT: Chlorini, John
; APPLICANT: Kotin, Robert M.
; APPLICANT: Sauer, Brian
; APPLICANT: Davidson, Elizabeth

```

```

: APPLICANT: Zabner, Joseph
: TITLE OF INVENTION: AAV VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
: FILE REFERENCE: 14014.032302
: CURRENT APPLICATION NUMBER: US/09/533,427
: CURRENT FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 5
: LENGTH: 588
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
: OTHER INFORMATION: synthetic construct
US-09-533-427-5

```

```

Query Match      57.3%; Score 1665; DB 2; Length 588;
Best Local Similarity 58.9%; Pred. No. 7,66-144;
Matches 318; Conservative 54; Mismatches 154; Indels 14; Gaps 5

OY      1 MASGGAGPAMADNEGADGVGNASGNMHCDSITWLDGRTVITSTRTWALTPTYNHLYKQISS 60
DB      57 MSAGGGGPLGDNNGADGVGNASGDHMCDSITWGGRRVYTKSTRITVLVSYNHNYREIKS 116
OY      61 ASTGASDNHNYFGSTPMGWGFDFNRFHSHSPRDMORLINNMGRRPRKLNPKLENIQVK 120
DB      117 GSVDSNANNAFGYSTPMGWGFDFNRFHSHSPRDMORLINNMGRRPRSLRKVINIQVK 176
OY      121 EVTTNDGCVTIANNLITSTVQVFSDEYQLPYLVS AHQGLPPRPADYFMIPOYGYLTLN 180
DB      177 EVTVQDSITTTIANNLITSTVQVFTDDDYQLPYVNGVGTGCLPAPFPQVFTLPQGYATLN 236
OY      181 --NSQAVGRSSPYCLLEYFSPSQMLRTGNPFSTFEVPRPHSSVAHSQSLDRLNPLID 238
DB      237 RDNTENPTERRSPFCLLEYFSPSKMLRTGNPFSTFEVPRPHSSVAPQNLKLANPLVD 296
OY      239 QYLYLNRTQNSGSAQKIDLFSRGSAPGMSVCP--KMLPGPCYCRQORVSKTKTDNNNS 297
DB      297 QYLRFSVTNNTGVOYQFNKL-----AGRYANTYKMPPOPMQRTQGMNLGSGVNRAS 349
OY      298 NFWTGSASKYNLNGRESIIINPGTAMASHKDEDEKFFPMSGVMI FGKESAGASNTAL--D 354
DB      350 VSAFATYTRMELGASGVQVPEQPMGMTNNLQGSNTYALENTMI FNSQPMNPGTTATYLEG 409
OY      355 NVMTIDEEIKATNPVATERFGTVAVNFOSSTDPATGDVHAMGALPGVWQDRDVTYLOG 414
DB      410 NMLTTSSETPVNRVAVNVAGQVAMYNQSSSTTAPATGTYINQELVPQSVMMERDVTYLOG 469
OY      415 PIWAKIPTTDGHFHPSPLMGFGGLKNPPQOLIKNTPVNPAPPAESATKFASTIYQYST 474
DB      470 PIMAKIPIPTGAHFHSPAMGGGLGHPPPMULIKNTPVGN--ITSFSDVPVSSFIYQYST 528
OY      475 GQVSVIEIMELQKENSRTWNPBEVQYTSINVASANDPTVDDNGLYTERPIGTRYITRPL 534
DB      529 GQVTEWEMELKKENSKRMNPEIQYTNVYNDPQVDFAPADSTGEYRTRTPIGTRYITRPL 588

RESULT 15
US-09-533-427-4
; Sequence 4, Application US/09533427
; Patent No. 6855314
; GENERAL INFORMATION:
; APPLICANT: Chlorini, John
; APPLICANT: Kotlin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Elizabeth
; APPLICANT: Zabrner, Joseph
; TITLE OF INVENTION: AAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LONG CELLS
; FILE REFERENCE: 14014, 0323U2
; CURRENT APPLICATION NUMBER: US/09/533,427
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
; OTHER INFORMATION: synthetic construct
US-09-533-427-4

```

```

Query Match      57.3%; Score 1665; DB 2; Length 724;
Beet Local Similarity 58.9%; Pred. No. 1e-143;
Matches 318; Conservative 54; Mismatches 154; Indels 14; Gaps 5;

```

```

QY 1 MASGGGAPADNNEGADGYNASGNMHCSTWLGDRVITTSRTWALPTYNHLYKOISS 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 193 MSAGGGGPGIDNQAADGYNASGDMHCSTWMDRVVTKSTRVWLPSTYHQRKIKS 252
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 ASTGASNDNHYFGYSTPMGYDPDFNRFCHFSFPRDQRLINNMGFRPKLNFKLFIQVK 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 253 GSVDSGNANAYFGYSTPMGYDPDFNRFCHFSFPRDQRLINNMGFRPKLNFKLFIQVK 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 EYTNNDGVTIANNTLSTYQVFSSEYQIPYVLSAHQCLPPPADYFMIPQYGLTLN 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 313 EYTVQDSTTTIANNTLSTYQVFTDDYQIPYVVGNGTEGCLPAPFPQVFTLPQYGYATLN 372
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 181 --NGSQAVGRSSFFCYCLEYFPPSQMLRTGNNFTFSYFEEVFPFHSYAHQSGLDRLNNPLID 238
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 373 RDNTEPTEPERSFFCYCLEYFPPSQMLRTGNNFEFTNFEEVFPFHSYAHQSGLDRLNNPLID 432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 239 QYLYVILNRTQNSGSAONKDLFSRGSPPAGMSVOP-KQMLPGPCYRQORVSKTKTDNNNS 297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 433 QYLYRFVSTNNNGVQVFNKUL-----AGRYANTYKQWFGPPMGRTQGMNLGSGVNRAS 485
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 298 NPTWTGASKYNNLNGRESIINPGTAMASHKDEDEKFFPMSCVMIKGSASAGASNTAL---D 354
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 486 VSAFATNTNMELEGASYQVPPQPNQGTNNLQGSNTYALENTMIFNSOPANPGTATYDEG 545
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 355 NMWITDEBEIKATNPVATERPGTVAVNFGSSSTDPAIGDVHAMGALPGWVMDRDVYLOG 414
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 546 NMLITSESETQPVNRYAVVVGQMATNNOSTTAPATGTYNLOETVPGSVMMERDVYLOG 605
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 415 PIWAKIPHTDGHFHPSPPLMGFGGLKNPPOLLIKNTVPVNPANPPASFSATKPFASFTQYST 474
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 606 PIWAKIPETGAHFPSPAMGFGGLKHPMMLIKNTVPVGN-ITSFSDVPVSSFTQYST 664
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 475 GQVSVIEIEMELQKNSKRNPPVQVTSNPAKSANDFTVDNNGLYTBPRPIGTIRYLTRPL 534
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 665 GQVTVEMEWELKNSKRNPPVQVTSNPAKSANDFTVDNNGLYTBPRPIGTIRYLTRPL 724
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Search completed: November 23, 2005, 17:42:49
 Job time : 25.2857 secs